

The Molecular Biology of Dehydration Tolerance: Regulation of Gene Expression and Function in *Xenopus laevis*

Liam James Hawkins

B.Sc. Hon. Dalhousie University, 2013

M.Sc. Carleton University, 2016

A thesis submitted to the Faculty of Graduate Studies and Research in partial
fulfilment of the requirements for the degree of

Doctor of Philosophy
Department of Biology

Carleton University,
Ottawa, Ontario, Canada

© Copyright 2020
Liam James Hawkins

The undersigned hereby recommend to the Faculty of Graduate Studies and Research
acceptance of this thesis

The Molecular Biology of Dehydration Tolerance: Regulation of Gene Expression and Function in *Xenopus laevis*

Submitted by
Liam James Hawkins, B.Sc. Hon., M.Sc.

in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Chair, Department of Biology

Thesis Supervisor

External Examiner

Carleton University

Abstract

The African clawed frog, *Xenopus laevis*, has been used as a model organism for cellular and developmental biology for nearly a century. Comparatively unstudied is its natural tolerance to dehydration brought about by seasonal drought evaporating its aquatic habitats. To survive the loss of >30% body water content, these animals employ several tissue-specific adaptations ranging from switching to ureotelism to relying on anaerobic metabolism as oxygen transport decreases with increased blood viscosity. Previous studies have indicated dehydration responsive gene expression and function is regulated with multiple mechanisms. In this thesis I further establish *X. laevis* as a dehydration tolerance model organism by determining suitable RT-qPCR reference genes in eight tissues. I then investigate regulatory mechanisms capable of large-scale regulation, namely, DNA methylation and histone modifications, microRNA, and reversible protein phosphorylation. Global levels of epigenetic marks showed little response to dehydration apart from increased 5hmC and decreased H3K4me in the liver, suggestive of epigenetic reprogramming. MicroRNAs, which are short RNAs that negatively regulate translation of specific mRNAs, were then examined in the heart. This analysis revealed a trend of downregulation during dehydration, and the enrichment of several important pathways including *cardiac muscle contraction* and *glycolysis and gluconeogenesis*. Particularly telling is the near uniform prediction of decreased regulation of all glycolytic enzyme transcripts that may support increased anaerobic glycolysis capacity during dehydration. Next, I analyzed the liver and skeletal muscle phosphoproteomes during dehydration and found a strong and concerted response by the liver and not muscle. Also emerging from the data was the significant upregulation and

phosphorylation of a hypoxia inducible PFKFB isozyme in the liver known to support glycolysis in many cancers. Together these results significantly advance our understanding of the molecular biology of dehydration tolerance and provide multiple clear directions for future studies.

Statement of Contribution

Except where specified, I designed and performed all experiments, analyzed the data, wrote all bioinformatic pipelines, created figures, and wrote all chapters. Dr. Kenneth B. Storey supported all chapters through funding, guiding this research program, editorial review, and final approval of all chapters. Janet M. Storey assisted all animal experiments and provided editorial review.

Chapter 1:

Excerpts from this chapter are adapted from:

Hawkins, L. J., & Storey, K. B. (2020). Advances and applications of environmental stress adaptation research. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, 240(February 2020), 110623. <https://doi.org/10.1016/j.cbpa.2019.110623>

The publisher, Elsevier, allows for inclusion of this article, in part or in full, in a thesis or dissertation as part of the rights of the author for personal use.

Chapter 2:

Bryan E. Luu assisted with some of the RNA extractions, cDNA synthesis, performed some of the RT-qPCR experiments, and approved the final manuscript.

This chapter has been published in *Gene Reports*:

Hawkins, L. J., Luu, B. E., & Storey, K. B. (2018). Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. *Gene Reports*, 13, 192–198. <https://doi.org/10.1016/j.genrep.2018.10.006>

The publisher, Elsevier, allows for inclusion of this article, in part or in full, in a thesis or dissertation as part of the rights of the author for personal use.

Chapter 4:

This study has been published in *Genomics*:

Hawkins, L. J. & Storey, K. B. (2020). MicroRNA expression in the heart of *Xenopus laevis* facilitates metabolic adaptation to dehydration. *Genomics*. <https://doi.org/10.1016/j.ygeno.2020.04.003>

The publisher, Elsevier, allows for inclusion of this article, in part or in full, in a thesis or dissertation as part of the rights of the author for personal use.

Chapter 5:

Xiaoshuang Wang, Xiaomin Xue, and Hui Wang performed the LC-MS/MS analysis, matched mass-spectra to putative proteins, and approved the final manuscript.

Acknowledgements

First, I want to thank my supervisor Dr. Kenneth Storey for bringing me, a student with no experience, into his lab and fostering an environment of growth that allowed me to become the person that I am today. Thank you for giving me the freedom to scrape my knees and for picking me back up when I did. I would also like to thank Jan Storey for being the backbone of the lab. Thank you for the countless hours spent molding me into a presentable writer and for always being there to give invaluable advice.

I want to thank all members of the Storey lab that I have worked with, particularly those who taught me invaluable skills and collaborated with me over the years. Sanoji Wijenayake and Kama Szereszewski for being handed a blank slate and teaching me fundamental laboratory skills, Bryan Luu for the long days of PCR made easy by great conversation, Rasha Al-attar and Alex Watts for commiserating with me as we started and completed our degrees together, Sam Logan for years being desk mates and sharing my desire for hard work punctuated with conversation, venting, and reflection, and Hanane Hadj-Moussa for being my scientific partner in crime and being there to sanity-check my crazed ideas.

Most importantly, I want to thank my parents for instilling in me the ambition to do what I want to do and to do it well. Everyday I see more and more characteristics of myself that are reflections of you two which gives me an immense sense of pride. Dad, thank you for showing me what hard work looks like and what it means to be a person who can always be relied on. Mom, thank you for always supporting and encouraging my various endeavours and teaching me not to strive for “good enough”. You two are the reason I am where I am today, and I cannot thank you enough.

Table of Contents

Title Page.....	i
Acceptance Sheet.....	ii
Abstract.....	iii
Statement of Contribution.....	v
Acknowledgements.....	vi
Table of Contents.....	vii
List of Abbreviations.....	xi
List of Tables.....	xv
List of Figures.....	xvii
List of Appendices.....	xxiii
1 Chapter 1: General Introduction.....	1
1.1 Introduction.....	2
1.2 Dehydration tolerance of <i>Xenopus laevis</i>	3
1.2.1 Physiological and metabolic adaptations.....	5
1.2.2 Enzymatic adaptations.....	6
1.2.3 Transcriptional and translational response to dehydration.....	9
1.3 The need for internal reference controls.....	10
1.4 Potential regulators of dehydration tolerance.....	12
1.4.1 DNA methylation.....	14
1.4.2 Histone modifications.....	16
1.4.3 MicroRNA.....	19
1.4.4 Reversible protein phosphorylation.....	21
1.5 Objectives.....	23
1.5.1 Objective 1.....	23
1.5.2 Objective 2.....	24
1.5.3 Objective 3.....	24
1.5.4 Objective 4.....	25
1.6 References.....	27
2 Chapter 2.....	36
2.1 Abstract.....	38
2.2 Introduction.....	39
2.3 Materials and Methods.....	41
2.3.1 Animal care and treatment.....	41
2.3.2 RNA extraction.....	42
2.3.3 cDNA synthesis.....	42
2.3.4 Candidate reference gene selection and primer design.....	43
2.3.5 RT-qPCR.....	43
2.3.6 Reference gene stability analysis and validation.....	44
2.3.7 Determination of optimal number of reference genes.....	45
2.4 Results.....	46
2.4.1 Expression levels of candidate reference genes.....	46
2.4.2 GeNorm Analysis.....	46
2.4.3 Normfinder Analysis.....	47

2.4.4	BestKeeper Analysis	47
2.4.5	Candidate reference gene composite ranking.....	48
2.4.6	Optimal number of reference genes for each tissue	48
2.5	Discussion	49
2.5.1	Candidate reference gene expression varies by tissue	50
2.5.2	Composite ranking show tissue-specific order.....	51
2.5.3	Two reference genes are sufficient in most tissues.....	52
2.5.4	Conclusions	52
2.6	References	54
3	Chapter 3	68
3.1	Abstract	70
3.2	Introduction	71
3.3	Materials and Methods	74
3.3.1	Animal treatment.....	74
3.3.2	Genomic DNA extraction.....	75
3.3.3	Quantification of DNA modifications.....	76
3.3.4	Nuclear fractionation and protein extraction.....	77
3.3.5	DNMT activity assay	77
3.3.6	RNA extraction and cDNA synthesis.....	78
3.3.7	Primer design and RT-qPCR.....	78
3.3.8	Western blotting	79
3.3.9	Statistics	80
3.4	Results	80
3.4.1	Global DNA modification levels.....	80
3.4.2	DNMT activity levels.....	81
3.4.3	<i>Dnmt</i> expression.....	81
3.4.4	DNA demethylation gene expression.....	81
3.4.5	Histone modifications	82
3.5	Discussion	82
3.5.1	DNA methylation status during dehydration	83
3.5.2	Hypoxia-inducible TETs highly upregulated in skeletal muscle	85
3.5.3	Histone methylation and acetylation largely unaffected by dehydration	86
3.5.4	Conclusions	87
3.6	References	89
4	Chapter 4	99
4.1	Abstract	101
4.2	Introduction	102
4.3	Materials and Methods	105
4.3.1	Animal experiments	105
4.3.2	RNA extraction	106
4.3.3	Small RNA sequencing.....	106
4.3.4	Read processing.....	107
4.3.5	Differential expression analysis and clustering.....	107
4.3.6	Gene set analysis	108
4.3.7	Three prime untranslated region analysis.....	109
4.3.8	Data visualization	110

4.3.9	MicroRNA cDNA synthesis.....	110
4.3.10	RT-qPCR Validation	111
4.4	Results	112
4.4.1	Small RNA sequencing summary	112
4.4.2	Differential expression of microRNA in response to dehydration.....	113
4.4.3	Gene ontology terms enriched for differentially expressed microRNA	114
4.4.4	KEGG pathways with reduced microRNA regulation.....	116
4.5	Discussion	117
4.5.1	Most microRNA are downregulated during dehydration.....	117
4.5.2	RNA and DNA binding targets highly enriched during dehydration ...	119
4.5.3	MicroRNAs targeting cardiac muscle contraction are downregulated .	122
4.5.4	Downregulation of microRNAs targeting most glycolytic enzymes	123
4.5.5	Conclusions	125
4.6	References	127
5	Chapter 5	159
5.1	Abstract	161
5.2	Introduction	162
5.3	Materials and Methods	165
5.3.1	Animal experiments	165
5.3.2	Protein extraction	166
5.3.3	Phosphoproteomic analysis	166
5.3.4	Data analysis and differential phosphorylation analysis	166
5.3.5	Hierarchical clustering and principle component analysis.....	167
5.3.6	Gene set analysis	167
5.3.7	RNA extraction and cDNA synthesis.....	168
5.3.8	Primer design and RT-qPCR.....	169
5.3.9	Data visualization.....	170
5.4	Results	171
5.4.1	Phosphoproteomics summary	171
5.4.2	Dehydration produces distinct phosphorylation patterns	171
5.4.3	Differential phosphorylation in response to dehydration.....	172
5.4.4	Overrepresentation of GO terms and KEGG pathways	173
5.4.5	Phosphorylation and expression of PFKFB enzymes	174
5.5	Discussion	177
5.5.1	Consistent phosphorylation pattern in liver but not skeletal muscle.....	177
5.5.2	Central metabolic pathways are differentially phosphorylated.....	179
5.5.3	Structural proteins only commonality between liver and muscle	180
5.5.4	Regulators of glycolysis phosphorylated during dehydration.....	181
5.5.5	Hypoxia-inducible PFKFB3 highly upregulated and phosphorylated ..	182
5.5.6	Conclusions	184
5.6	References	185
6	Chapter 6: General Discussion	204
6.1	Discussion	205
6.2	Reference gene determination	206
6.2.1	Evaluation of objective 1.....	207
6.3	Profile of epigenetic mechanisms in response to dehydration	208

6.3.1	Evaluation of objective 2.....	210
6.4	Differential expression of microRNA during dehydration	210
6.4.1	Evaluation of objective 3.....	214
6.5	Phosphoproteomics reveals <i>pfkfb3</i> phosphorylation and expression	214
6.5.1	Evaluation of objective 4.....	218
6.6	Future directions	218
6.7	Final conclusions	222
6.8	References	224
Appendices		228
Appendix A: List of publications		229
Appendix B: Communications at scientific meetings		230
Appendix C: RT-qPCR quantification		231
Appendix D: Full lane Western blot images		235
Appendix E: Methylation-sensitive high-resolution melt analysis		236

List of Abbreviations

3PO - 3-(3-pyridinyl)-1-(4-pyridinyl)-2-propen-1-one	221
5caC - 5-Carboxylcytosine	15
5fC - 5-Formylcytosine.....	15
5hmC - 5-Hydroxymethylcytosine.....	15
5mC - 5-Methylcytosine.....	14
ACTB - Beta actin	40
AGO - Argonaut protein.....	20
ALDOB - Fructose-bisphosphate aldolase.....	174
ATP - Adenosine triphosphate	7
ATP5BF1 - ATP synthase F1 subunit beta	40
BLAST - Basic local alignment search tool	43
bp - Base pair.....	43
BWC_i - Initial body water content	41
CANX - Calnexin	40
Cas - CRISPR associated protein	16
cDNA - Complementary DNA	10
ChIP - Chromatin-immunoprecipitation	219
CoA - Coenzyme A	17
C_q - Quantification cycle.....	44
C_{qcorr} - Corrected quantification cycle	45
C_{qmin} - Minimum quantification cycle	44
C_{qsample} - Sample quantification cycle.....	44
CRISPR - Clustered regularly interspaced short palindromic repeats.....	220
CS - Citrate synthase	174
dCas9 - Endonuclease dead Cas9.....	16
ddH₂O - Double distilled H ₂ O	42
DGCR - DiGeorge syndrome chromosomal region.....	20
DMR - Differentially modified regions.....	220
DMSO - Dimethyl sulfoxide	43
DNMT - DNA methyltransferase.....	14
dNTP - Deoxyribose nucleoside triphosphate.....	43
dsRNA - Double stranded RNA.....	20
dT - Deoxythymidine	42
DTT - Dithiothreitol	77
E% - Efficiency percentage	44
EDTA - Ethylenediaminetetraacetic acid.....	75
EGTA - Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid.....	166
ELISA - Enzyme-linked immunosorbent assay	76
ENO - Enolase.....	116
ERK - Extracellular signal-regulated kinase.....	9
FBP2 - Fructose-2,6-bisphosphatase	175
FC - Fold change	107
FDR - False discovery rate.....	107
FOXO - Forkhead box protein O	9

Fru-2,6-P₂ - Fructose 2,6-bisphosphate	175
Fru-6-P - Fructose 6-phosphate	175
GAPDH - Glyceraldehyde 3-phosphate dehydrogenase	11
gDNA - Genomic DNA	237
GMT - Gene matrix transposed.....	168
GO - Gene ontology	108
GSK3β - glycogen synthase kinase-3 beta	174
GST - Glutathione S-transferase	9
HAT - histone acetyltransferase	17
HDAC - Histone deacetylase.....	16
HEPES - 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid	77
HIF - Hypoxia inducible factor	86
HK - Hexokinase	116
HMT - Histone methyltransferase.....	17
HRE - Hypoxia-responsive element.....	176
HRP - Horseradish peroxidase	76
HSC - Heat shock cognate protein	20
HSP - Heat shock protein	20
KC7F2 - N,N'-(Dithiodi-2,1-ethanediyl)bis[2,5-dichloro-benzenesulfonamide.....	222
kDa - Kilodalton.....	217
KEGG - Kyoto Encyclopedia of Genes and Genomes	108
K_m - Michaelis-Menten Constant.....	7
LC-MS/MS - Liquid chromatography - tandem mass spectrometry	164
LDH - Lactate dehydrogenase.....	7
M - Stability value	46
MAPK - Mitogen-activated protein kinase	20
MBD - Methyl-binding domain.....	15
m_d - Dehydrate mass	41
MeCP - Methyl-CpG binding protein	15
m_i - Initial mass.....	41
M-MLV - Moloney murine leukemia virus	43
MnSOD - Manganese-dependent superoxide dismutase	10
mRNA - Messenger RNA	10
MS-HRMA - Methylation-sensitive high-resolution melt analysis.....	219
NAD⁺ - Nicotinamide adenine dinucleotide	7
NF - Normalization factor	45
NFκB - Nuclear factor kappa-light-chain-enhancer of activated B cells	173
NGS - Next-generation sequencing	4
NRF - Nuclear factor erythroid 2-related factor.....	9
OD - Optical density.....	42
PCA - Principle component analysis.....	167
P_{CO2} - Partial pressure of carbon dioxide.....	102
PCR - Polymerase chain reaction.....	43
PEP - Phosphoenolpyruvate.....	7
PFK1 - Phosphofructokinase 1.....	116
PFK15 - 1-(4-pyridinyl)-3-(2-quinolinyl)-2-propen-1-one	221

PFK2 - Phosphofructokinase-2	175
PFKFB - 6-Phosphofructo-2-kinase/fructose-2,6-biphosphatase	174
PGK - Phosphoglycerate kinase	213
PGM - Phosphoglucomutase	174
piRNA - PIWI interacting RNA	107
PK - Pyruvate kinase	116
pre-microRNA - Precursor microRNA	20
pri-microRNA - Primary microRNA	20
PVDF - Polyvinylidene difluoride.....	79
Q - Relative quantity	44
RanGTP - GTP-binding nuclear protein Ran	20
REVIGO - Reduce and visualize gene ontology	109
RISC - RNA-induced silencing complex	20
RNA - Ribonucleic acid.....	20
RNA Pol II - RNA polymerase II	19
ROI - Region of interest	236
RPL27 - Ribosomal protein L27	40
rRNA - Ribosomal RNA	107
RSK - Ribosomal S6 kinase	9
RT-PCR - Reverse-transcription polymerase chain reaction.....	10
RT-qPCR - Reverse-transcription quantitative polymerase chain reaction.....	10
RyR - Ryanodine receptor	123
SAH - S-adenosyl-homocysteine.....	14
SAM - S-adenosyl-methionine	14
SD - Standard deviation	45
SDHA - Succinate dehydrogenase complex subunit A	40
SDS - Sodium dodecyl sulfate	77
SDS-PAGE - Sodium dodecyl sulfate-polyacrylamide gel electrophoresis	79
SEM - Standard error of the mean.....	80
SERCA - Sarco/endoplasmic reticulum Ca ²⁺ -ATPase.....	123
S_{microRNA} - MicroRNA score	108
S_{mRNA} - mRNA score	108
snoRNA - Small nucleolar RNA	107
snRNA - Small nuclear RNA	107
SRA - Sequence Read Archive.....	107
ssRNA - Single stranded RNA	20
STAT - Signal transducer and activator of transcription.....	9
TBP - TATA-box binding protein.....	40
TBST - Tris buffered saline with Tween-20	79
TCA - Tricarboxylic acid	8
TDG - Thymine DNA glycosylase.....	15
TET - Ten-eleven translocation.....	15
TPI - Triosephosphate isomerase	117
TPM - Tropomyosin.....	122
tRNA - Transfer RNA	107
TSS - Transcription start site	176

UPLC - Ultra performance liquid chromatography	166
UTR - Untranslated region	19

List of Tables

Table 2.1 Candidate reference gene primers used for qPCR analysis in <i>X. laevis</i> . Each primer pair was used for all eight tissues of interest. NCBI RefSeq accession numbers are provided when available. The Xenbase accession number is provided for <i>atp5f1b.L</i>	57
Table 2.2 Composite rankings for reference gene candidates in eight tissue of <i>X. laevis</i> . Candidate reference genes are first ranked by stability using geNorm, Normfinder, and BestKeeper software. These ranks are averaged for each candidate gene in each tissue to get a composite rank. Composite ranks are in parenthesis and top <i>N</i> genes as determined in Figure 2.4 are bold.....	58
Table 3.1 Primers used for RT-qPCR analysis of <i>Xenopus laevis</i> . Both short (S) and long (L) chromosome sub-genome genes of each target were measured unless one version did not exist. NCBI RefSeq or nucleotide accession numbers are provided for each gene. Reference genes for normalization (<i>atp5f1b.L</i> and <i>sdha.L</i> for liver, <i>rpl27.L</i> and <i>gapdh.S</i> for skeletal muscle) were used as previously determined (Hawkins et al. 2018b).	92
Table 5.1 Glucose metabolism related enzymes corresponding to identified phosphopeptides. Glycogen metabolism, glycolysis, gluconeogenesis, and TCA cycle enzymes with identified phosphopeptides are presented. FDR corrected <i>p</i> -values in bold are statistically significant ($p < 0.05$).	189
Supplementary Table 2.1 C_q values for brain, heart, muscle, and liver of 0%, 15%, and 30% dehydrated <i>Xenopus laevis</i> (mean \pm SE; $n=4$ per condition), and kidney, dorsal skin, ventral skin, and lung of 0% and 30% dehydrated <i>X. laevis</i> (mean \pm SE; $n=6$ per condition)	63
Supplementary Table 2.2 GeNorm candidate reference gene stability rankings for eight tissues from dehydrated <i>Xenopus laevis</i> . Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Stability (M) values are in parenthesis	64
Supplementary Table 2.3 Normfinder candidate reference gene stability rankings for eight tissues from dehydrated <i>Xenopus laevis</i> . Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Stability values are in parenthesis	65
Supplementary Table 2.4 BestKeeper candidate reference gene stability rankings for eight tissues from dehydrated <i>Xenopus laevis</i> . Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Standard deviation values are in parenthesis	66
Supplementary Table 4.1 Primers used for RT-qPCR validation of microRNA-seq ..	140
Supplementary Table 4.2 MicroRNA-Seq summary and raw read processing results	141
Supplementary Table 4.3 Differential expression results of microRNAs in heart of <i>Xenopus laevis</i> exposed to dehydration (Data available in: Hawkins 2020)	142

Supplementary Table 4.4 GO Molecular Functions enriched by microRNAs differentially expressed in the heart of <i>Xenopus laevis</i> during dehydration (Data available in: Hawkins 2020)	143
Supplementary Table 4.5 GO Biological Processes enriched by microRNAs differentially expressed in the heart of <i>Xenopus laevis</i> during dehydration (Data available in: Hawkins 2020)	144
Supplementary Table 4.6 GO Cellular Components enriched by microRNAs differentially expressed in the heart of <i>Xenopus laevis</i> during dehydration (Data available in: Hawkins 2020)	145
Supplementary Table 4.7 KEGG pathways enriched by microRNAs differentially expressed in the heart of <i>Xenopus laevis</i> during dehydration (Data available in: Hawkins 2020)	146
Supplementary Table 4.8 Predicted MicroRNA:mRNA target pairing of microRNAs expressed in the heart of <i>Xenopus laevis</i> in response to dehydration (Data available in: Hawkins 2020).....	147
Supplementary Table 4.9 MicroRNAs expressed in the heart of <i>Xenopus laevis</i> during dehydration that target genes in the KEGG Cardiac Muscle Contraction pathway (Data available in: Hawkins 2020)	148
Supplementary Table 4.10 MicroRNAs expressed in the heart of <i>Xenopus laevis</i> during dehydration that target genes in the KEGG Glycolysis Gluconeogenesis pathway (Data available in: Hawkins 2020)	149
Supplementary Table 5.1 Primer sequences for qPCR analysis.	199

List of Figures

- Figure 1.1** Phylogenetic tree of select model and dehydration tolerant animals. Species in blue are dehydration tolerant. Species with a black arrow have well-annotated genomes. *Xenopus laevis* is the most closely related species to humans that is dehydration tolerant (blue) and has a fully annotated genome (arrow). The phylogenetic tree was generated using Interactive Tree of Life (iTOL) (Letunic and Bork 2016). 34
- Figure 1.2** Select mechanisms of regulation potentially contributing to dehydration tolerance in *Xenopus laevis*. As the environment becomes arid and dehydration progresses, signals are transduced to a variety of regulatory mechanisms that enact adaptational responses enabling survival until rehydration occurs. The mechanisms of regulation studied here start at the epigenetic level where DNA methylation and histone modifications are examined, then an analysis of the microRNAome is undertaken, followed by phosphoproteomic analysis. 35
- Figure 2.1** Distribution of quantification cycles of all samples for each candidate reference gene. Boxes extend from the first to third quartile where the center horizontal line is the second quartile. Whiskers extend to the 5th and 95th percentiles 59
- Figure 2.2** Mean quantification cycle for each experimental condition of each tissue. Data are mean \pm SE, n=4-6 depending on tissue 60
- Figure 2.3** Composite and component rankings for reference gene candidates in eight tissue of *X. laevis*. Lower ranks indicate better stability and suitability as a reference gene. Component stabilities are produced by (\blacklozenge) geNorm, (\blacksquare) Normfinder, (\blacktriangle) BestKeeper. The composite ranking (\bullet) is the mean of all component rankings..... 61
- Figure 2.4** Determination of optimal number of reference genes. Pairwise variation (V) of n and $n+1$ gene containing normalization factors (NF_n, NF_{n+1}). Higher V indicate higher effect of adding the next most stable reference gene by composite rank (Table 2.2) for normalization. The recommend cut-off of 0.15 from (Vandesompele et al. 2002) was used to determine when a subsequent reference gene should be included. The optimal number is indicated for each tissue 62
- Figure 3.1** Global DNA modification levels in response to dehydration in *Xenopus laevis*. Global levels of a) 5-methylcytosine (5mC) and b) 5-hydroxymethylcytosine (5hmC) were measured in liver and skeletal muscle tissues. Values are mean \pm SEM of modification as percentage of genomic DNA, asterisks indicate statistical significance ($p < 0.05$) using a Student's t-test..... 93
- Figure 3.2** DNA methyltransferase activity in response to dehydration in *Xenopus laevis*. Values are mean \pm SEM, asterisks indicate statistical significance ($p < 0.05$) using a Student's t-test. 94
- Figure 3.3** Relative expression of *dnmt* genes in *Xenopus laevis* during dehydration. *Dnmt* genes were measured by RT-qPCR in a) liver and b) skeletal muscle. Short chromosome (S) sub-genome gene of each target are relative to the control levels of their respective long chromosome (L) sub-genome gene. Values are mean \pm SEM

and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.....	95
Figure 3.4 Relative expression of DNA demethylation genes in <i>Xenopus laevis</i> during dehydration. DNA demethylation genes were measured by RT-qPCR in a) liver and b) skeletal muscle. Short chromosome (S) sub-genome gene of each target are relative to the control levels of their respective long chromosome (L) sub-genome gene. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.....	96
Figure 3.5 Relative changes in histone modifications in the liver of <i>Xenopus laevis</i> in response to dehydration. Global levels of histone modifications were measured in the liver using western blotting. All modification levels are relative to total levels of histone H3. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.	97
Figure 3.6 Relative changes in histone modifications in the skeletal muscle of <i>Xenopus laevis</i> in response to dehydration. Global levels of histone modifications were measured in skeletal muscle using western blotting. All modification levels are relative to total levels of histone H3. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.	98
Figure 4.1 MicroRNA differentially expressed in the heart of <i>Xenopus laevis</i> in response to dehydration. Volcano plot fold-change thresholds were set to $\pm \log_{21.5}$ and an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate microRNA that were significantly downregulated and upregulated, respectively and grey circles are microRNA that did not pass the fold-change and p -value thresholds.....	132
Figure 4.2 Hierarchical clustering of significant differentially expressed microRNA in the heart of <i>Xenopus laevis</i> in response to dehydration. Clustering of samples and microRNA was done using the Ward method. Purple and orange indicate higher and lower z-scores (standardized expression level) for each microRNA in each sample. Clustering of samples at the highest level of the dendrogram is indicated by dark and light blue bar above the heat map	133
Figure 4.3 Enrichment of GO Molecular Function terms in heart of dehydrated <i>Xenopus laevis</i> . a) Volcano plot of GO Molecular Function term enrichment with an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. b) Top 15 most significantly enriched (by FDR-adjusted p -value) GO Molecular Function terms	134
Figure 4.4 Enriched GO Molecular Function semantic relation network in heart of dehydrated <i>Xenopus laevis</i> . Enriched GO Molecular Function terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation	

network of summary terms was plotted using Cytoscape. The largest densely connected network comprised two sub-networks related to DNA and RNA binding (a) and transcription factor binding (b). Three small networks related to transporter and channel activities (c), post-translational modification activities (d), and molecular cleavage activities (e). Node size is proportional to the number of genes each term encompasses and the darker the blue color of a node is the lower the FDR-adjusted p -value and grey edges indicate a semantic relationship..... 136

Figure 4.5 Enrichment of KEGG pathways in heart of dehydrated *Xenopus laevis*. a) Volcano plot of KEGG pathway enrichment. b) All significantly enriched KEGG pathways. All other information as in Figure 4.3 137

Figure 4.6 MicroRNA that target genes in the Cardiac Muscle Contraction KEGG pathway. Genes and microRNA are colored by their *SmRNA* and *SmicroRNA* scores respectively, where a darker blue color indicates a more positive score and a darker red color indicates a more negative score. *SmRNA* is equal to the negative of the summation of *SmicroRNA* scores of all microRNA that target that mRNA. For more information about *SmRNA* and *SmicroRNA* scores see 4.5 Materials and Methods for definitions. A full list of genes encompassed by each label is found in Supplementary Table 4.9 138

Figure 4.7 MicroRNA that target genes in the Glycolysis Gluconeogenesis KEGG pathway. All other information as in Figure 4.6. A full list of genes encompassed by each label is found in Supplementary Table 4.10 139

Figure 5.1 Hierarchical clustering analysis of phosphopeptides in the liver and muscle of *Xenopus laevis* in response to dehydration. Clustering was performed on samples from control and dehydrated animals in a) liver and b) skeletal muscle. Samples (arranged horizontally) were clustered by Euclidean distance between relative quantities of phosphopeptides, and phosphopeptides (arranged vertically) were clustered based on Euclidean distance between each sample. Colors indicate Z-score (standardized quantities) where red is a high Z-score and blue is a low Z-score. The experimental group the sample belongs to is indicated by blue (control) or red (dehydrated) bars above the heatmap. 190

Figure 5.2 Principal component analysis of samples in the liver and muscle of *Xenopus laevis* in response to dehydration. Principle component analysis was performed on phosphopeptide relative quantities in a) liver, b) skeletal muscle, and c) both tissues. The cumulative explained variance of principal components is plotted in d). 191

Figure 5.3 Volcano plot of phosphopeptides from liver and muscle of *Xenopus laevis* in response to dehydration. Phosphopeptides that had significantly increased (FDR-corrected p -value < 0.05) and decreased relative abundance in dehydration samples are in blue and red respectively for a) liver and b) skeletal muscle. Phosphopeptides from liver and skeletal muscle ordered by Log₂ fold-change is shown in c). 192

Figure 5.4 Venn diagram of phosphoproteins from liver and muscle of *Xenopus laevis* in response to dehydration. Phosphoproteins with at least one corresponding phosphopeptide identified in liver and/or muscle is shown in a), whereas only those

with at least one significantly differentially abundant phosphopeptides are shown in b).	193
Figure 5.5 Enriched Gene Ontology biological processes semantic relation network in the liver of <i>Xenopus laevis</i> exposed to dehydration. Enriched GO biological processes terms were summarized by REVIGO (Supek et al. 2011) and resulting semantic relation network was plotted using Cytoscape. Node size is proportional to the number of genes each term encompasses and the darker the blue the lower the FDR-adjusted <i>p</i> -value. Edges indicate semantic relationship.....	194
Figure 5.6 Enriched Gene Ontology term semantic relation networks in skeletal muscle of <i>Xenopus laevis</i> exposed to dehydration. GO biological processes are in a), cellular compartments in b), and molecular functions in c). All other details as in Figure 5.5.	195
Figure 5.7 Top ten enriched Kyoto Encyclopedia of Genes and Genomes pathways in the liver of <i>Xenopus laevis</i> exposed to dehydration. Terms in dark blue are statistically enriched (FDR-corrected <i>p</i> -value < 0.05), terms in light blue are not significantly enriched.....	196
Figure 5.8 Phosphorylation of liver PFKFB enzymes in response to dehydration. PFKFB1 S31, PFKFB1 S32, and PFKFB3 S463 were identified as differentially phosphorylated in the liver during dehydration and are shown in a). S31 and S32 are present in the PFK-2 domain (translucent orange) of PFKFB1 as shown in b). S463 is present in the FBP-2 domain (translucent green) of PFKFB3 as shown in c). Arrows in b) and c) show phosphorylated serines (red) and substrates (blue).	197
Figure 5.9 Expression of <i>pfkfb3</i> and <i>pfkfb1</i> in the liver during dehydration. Primers were designed to measure the expression of <i>pfkfb3</i> transcript variants as shown by colored arrows in a). Blue arrows show the position of forward and reverse primers for measuring all four <i>pfkfb3</i> transcript variants, whereas red arrows show primer pairs for measuring specific variants. Primer sequences are in Supplementary Figure 5.1. b) shows relative expression of <i>pfkfb3</i> transcript variants in the liver of control and dehydrated <i>X. laevis</i> , whereas relative expression of <i>pfkfb1</i> is in c). d) shows the position of hypoxia responsive elements (HREs, binding sites for HIF1 α) within 2 kbp upstream of the transcription start sites (TSS) of <i>pfkfb1</i> and <i>pfkfb3</i>	198
Supplementary Figure 2.1 Reference gene qPCR primer validation. Melting curve derivative plots show a single prominent peak for each gene indicating a single product was amplified. Analysis using a 3% agarose gel (bottom right) show PCR products at expected sizes	67
Supplementary Figure 4.1 Genomic location of significantly differentially expressed microRNAs. Mixed indicates the microRNA is located in the genome where multiple genes overlap.	150
Supplementary Figure 4.2 RT-qPCR Validation results of differentially expressed microRNA.....	151

Supplementary Figure 4.3 Enrichment of GO Biological Processes terms in heart of dehydrated *Xenopus laevis*. a) Volcano plot of GO Biological Processes term enrichment with an FDR-adjusted *p*-value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. b) Top 15 most significantly enriched (by FDR-adjusted *p*-value) GO Biological Processes terms..... 152

Supplementary Figure 4.4 Enriched GO Biological Processes semantic relation network in heart of dehydrated *Xenopus laevis*. Enriched GO Biological Processes terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation network of summary terms was plotted using Cytoscape 154

Supplementary Figure 4.5 Enrichment of GO Cellular Components terms in heart of dehydrated *Xenopus laevis*. a) Volcano plot of GO Cellular Components term enrichment with an FDR-adjusted *p*-value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. b) Top 15 most significantly enriched (by FDR-adjusted *p*-value) GO Cellular Components terms..... 155

Supplementary Figure 4.6 Enriched GO Cellular Components semantic relation network in heart of dehydrated *Xenopus laevis*. Enriched GO Cellular Components terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation network of summary terms was plotted using Cytoscape 157

Supplementary Figure 4.7 Comparison human and *Xenopus laevis* microRNAs targeting of 3'UTRs of glycolytic genes. a) STarMir (<http://sfold.wadsworth.org/cgi-bin/starmirtest2.pl>) was used to determine logistic probabilities of microRNAs targeting 3'UTRs of human and *X. laevis* genes involved in glycolysis and gluconeogenesis from Figure 4.7. Statistical differences were determined with Student's t-test where * indicates *p*-value < 0.05, ** *p*-value < 0.01, and *** *p*-value < 0.001. b) Maximum logistic probability for each gene. Black arrows indicate genes where *X. laevis* had a higher maximum logistic probability than humans ... 158

Supplementary Figure 5.1 Summary of phosphopeptide from liver and muscle. a) Number of phosphopeptides containing phosphoserine (pS), phosphothreonine (pT), and phosphotyrosine (pY) residues. b) number of phosphopeptides containing one, two, or three phosphorylated residues. 200

Supplementary Figure 5.2 Hierarchical clustering of significantly differentially abundant phosphopeptides from liver and muscle of *Xenopus laevis* exposed to dehydration. Clustering of significantly differentially abundance phosphopeptides from a) liver and b) muscle..... 201

Supplementary Figure 5.3 Semantic relation network of enriched GO cellular compartment terms in the liver of dehydrated <i>Xenopus laevis</i> . All other information as in Figure 5.5.....	202
Supplementary Figure 5.4 Semantic relation network of enriched GO molecular function terms in the liver of dehydrated <i>Xenopus laevis</i> . All other information as in Figure 5.5.	203
Appendix Figure 1 Sample schema for quantification of mRNAs and microRNAs using RT-qPCR.	232
Appendix Figure 2 Example RT-qPCR results. a) Amplification of 2-fold serial dilution standard curve (rainbow curves) shows expected equidistant C_q values (red line) and NTC samples amplifying much later than standard curve samples or not at all. b) Plot of log of the dilution factor versus C_q shows reaction efficiency of 95.0% and R^2 of 0.999. c) Overlay of amplification of samples of interest. d) Samples of interest fall near middle of standard curve as expected from dilution. e) Melt curve analysis indicates a single product was amplified in all samples.	234
Appendix Figure 3 Full lane Western blot images for a) Figure 3.5 and b) Figure 3.6. Each lane corresponds to the first control sample of their respective tissue. Since all Western blots were targeting histone H3 or modified histone H3 residues all lanes are aligned at 17 kDa as indicated by the arrow	235
Appendix Figure 4 Methylation standard curve proof of concept for <i>hif1a.L</i> TSS CpG island. a) The region of interest (ROI) was identified near the <i>hif1a.L</i> TSS. b) The ROI was amplified and the product was purified. The purified product was then split into two samples, one being <i>in vitro</i> methylated by SssI CpG methyltransferase (100% methylation) and the other being mock treated (0%). c) The 0% (blue) and 100% (red) methylated products are then amplified after serial dilution using the same methylation naïve primers to verify no amplification bias, then d) verified to have different melt peaks but e) are still the same size. Results in c), d), and e) indicate that the 0% and 100% methylated products only differ by sequence introduced by the <i>in vitro</i> methylation since they are amplified by the same primers and are the same size.	239
Appendix Figure 5 Full methylation standard curve for <i>hif1a.L</i> TSS CpG island. a) A range of methylation percentages is produced by mixing the 0% and 100% methylated product in different proportions. b) The melt profile is normalized using <i>uAnalyze</i> . c) The change in helicity at a specific temperature is chosen for the standard curve.	241

List of Appendices

Appendix A: List of publications	229
Appendix B: Communications at scientific meetings	230
Appendix C: RT-qPCR quantification	231
Appendix D: Full lane Western blot images	235
Appendix E: Methylation-sensitive high-resolution melt analysis	236

Chapter 1: General Introduction

1.1 Introduction

The range of environments that animals can survive is much larger than those that humans can. We live in a relatively narrow slice of conditions with specific temperatures, oxygen levels, food, and water availabilities. Many animals live well outside these tight ranges and have evolved specialized adaptations suited to each niche. For example, mammals such as ground squirrels and bears hibernate to survive the cold winters, while lower vertebrates like the red-eared slider turtle drop their metabolic rate to survive for months without oxygen at the bottom of frozen ponds. Studying these animals is important not only to determine the entire spectrum of conditions that complex life can endure, but also to take what we can learn and apply it for our benefit. Tissue preservation, ischemia-reperfusion damage, anoxia and freezing damage mitigation, and extreme dehydration are all examples of current biomedical problems that nature has solved and thus can provide us with steppingstones to learn from.

All adaptations to extreme environmental conditions involve coordinated regulation of behavioral, physiological, cellular, and biochemical processes with intrinsic and environmental signals incorporated at every stage. The functional components (enzymes) that carry out the chemical processes of life are post-translationally modified and allosterically regulated to adjust their activities in response to signal inputs. Above this, post-transcriptional regulators, such as microRNA that control which gene transcripts are translated, are also at play. Other mechanisms that mediate environmental stress-induced gene expression have been extensively reviewed including modulation of transcription factors, and changing epigenetic programs (Storey and Storey 2004; Storey 2015). A common finding in studies of animals adapted to environmental stresses is that what differentiates them from closely related species (or humans) that cannot survive these

same conditions, isn't the presence of novel genes, but rather specialized regulatory programs of genes and proteins that we largely share in common. This gives further credence for the application of findings from these animals to relevant biomedical or industrial issues that we currently face.

Liquid water is essential for life as we know it. Most organisms live where liquid water is ubiquitous. Even in biomes, such as deserts where water is sparse, life revolves around where the water is concentrated and prioritizes functions that mitigate water loss. Interestingly, many organisms have embraced niches where water is periodically absent. The best-known examples are small invertebrates such as tardigrades (Boothby et al. 2017), nematodes (Erkut et al. 2011), and mites (Worland and Block 1986) that can live for months, years or decades in anhydrobiotic states, but comparatively little is known about dehydration tolerance in vertebrate models.

1.2 Dehydration tolerance of *Xenopus laevis*

The African clawed frog, *Xenopus laevis*, has long been a model system in cell biology, developmental biology, and toxicology. *X. laevis* is an allotetraploid resulting from a hybridization event occurring between two diploid progenitor species 17 to 18 million years ago (Session et al. 2016). This species has two subgenomes, each with nine homologous chromosomes, one a *long* version (L) containing 13,781 protein coding genes and one a *short* version (S) containing 10,241 protein coding genes (Session et al. 2016). Most *X. laevis* researchers choose this model for pragmatic reasons such as its relatively cheap cost, short development time, large egg size, and accessibility at all developmental stages, but also because it is a stepping stone from mammalian models to distantly related zebrafish (Moody et al. 2016). While the *X. laevis* model has been

fruitful for the above-mentioned fields, only a small amount of research has been devoted to its adaptations in its natural environment, particularly its dehydration tolerance. *X. laevis* is a mostly aquatic anuran originating from southern Africa. They are found in ponds, lakes, and wetlands, and encounter seasonal drying of their habitats. In response to habitat desiccation, like other aestivating species, they either burrow into a layer of mud to minimize dehydration or attempt to move to a new pond (Loveridge 1976). Despite these behaviors, *X. laevis* can lose as much as 30% of its total body water until rainfall and rehydration of the pond (Hillman 1978a; Malik and Storey 2009a). This remarkable ability extends the applicability to biological research of an already versatile model organism.

Xenopus laevis is the ideal model for studying natural dehydration tolerance. Of all dehydration tolerant organisms, *Xenopus* is the only one that satisfies the following criteria: 1) traditional model organism characteristics (low cost, short generation time, easy handling/housing/ husbandry, species-specific resources), 2) extensive genomic characterization, and 3) phylogenetic proximity to mammals. As previously mentioned, *X. laevis* has long been a model organism in various biological disciplines, and thus has been modelled anatomically, physiologically, and developmentally under laboratory settings. With the falling costs for next-generation sequencing (NGS), the *X. laevis* genome has been updated with deeper reads, extensive annotations, and staged developmental and adult tissue transcriptomic profiles (Session et al. 2016). All of this information is aggregated into the Xenbase Model Organism Database for ease of access and community organization (Karimi et al. 2018). Finally, amphibians (including *X. laevis*) are the most closely related dehydration tolerant group to mammals (**Figure 1.1**).

With regard to studying dehydration tolerance, species closely related to humans hold a greater chance of providing translational information for biomedical issues. Hence, *X. laevis* is a promising tool to help answer dehydration related questions.

1.2.1 Physiological and metabolic adaptations

Dehydration tolerance is facilitated by multiple behavioral, physiological, metabolic, and gene regulatory factors. Firstly, as *X. laevis* are ammonotelic like other aquatic animals, when water is not present, they are at risk of building up toxic levels of ammonia. This is combated by turning on the urea cycle and storing nitrogenous waste products in the form of urea, a phenomenon shown to occur both in naturally estivating *X. laevis* and those exposed to evaporative dehydration in laboratory settings (Balinsky et al. 1961, 1967). The accumulation of urea at a faster rate than expected indicates that it may be serving more functions than just nitrogen storage, such as increasing osmolyte concentrations to limit water loss across the otherwise permeable skin of this animal (Hillman 1978b).

As *X. laevis* dehydrate, plasma volume decreases linearly, presenting a stress on the organism by limiting transport of gases, nutrients and metabolic wastes (Hillman 1978b). This shows an inverse relationship to hematocrit and resting heart rate which both increase with dehydration (Hillman 1978a). The increased heart rate is suggested to offset diminished oxygen transport capabilities by the cardiovascular system as plasma viscosity increases, leading to generalized hypoxia at higher levels of dehydration as indicated by rising whole body levels of lactate, an anaerobic end product (Hillman 1978a).

Urea may also modulate plasma oxygen transport. A study by Jokumsen and Weber (1980) indicated that the elevated urea levels were correlated with higher hemoglobin affinity for oxygen even when other osmolytes known to decrease affinity were also elevated. These authors suggested that under the long-term estivation experiments, increased affinity would limit oxygen delivery and therefore limit oxygen consumption at the mitochondria, promoting a lower metabolic rate. This could then be quickly reversed upon rehydration when urea is quickly excreted from the body, restoring normal hemoglobin affinity, oxygen transport, and metabolic rate.

Interestingly, systemic blood flow distribution is reprioritized during dehydration (Hillman and Sommerfeldt 1981). Using intraventricular injections of radionuclide-labelled microspheres, blood flow was monitored during dehydration. The results indicated that while blood flow distribution in most tissues remains constant, blood flow increases to the brain and a proportional decrease is seen in the hind limb skeletal muscle. The implications here are obvious given the importance of maintaining oxygen delivery to the brain to maintain neural integrity during dehydration at the expense of otherwise dormant skeletal muscle.

1.2.2 Enzymatic adaptations

Studies of animals that survive environmental stressors show that the majority of adaptations employed do not rely on novel genes specific to the animal. Rather, they use the genes, pathways, and metabolism common to most animals with nuanced differences to their regulation, function, and response to environmental conditions. This is particularly true for enzymes involved in central metabolic pathways where relatively

small changes in the primary structure of these enzymes may allow non-canonical regulation of metabolism.

Given the importance of anaerobic metabolism during dehydration, glycolytic enzymes have been examined for potential dehydration tolerance regulatory adaptations. The liver plays an active metabolic role during dehydration as it is producing high levels of urea through the urea cycle, an energetically costly process requiring three ATP molecules. As dehydration increases and oxygen transport decreases, the urea cycle is presumably fueled by anaerobic glycolysis as indicated by increased lactate levels. The last two steps of anaerobic glycolysis are catalyzed by pyruvate kinase and lactate dehydrogenase (LDH) and are responsible for producing ATP and regenerating NAD^+ needed for further rounds of glycolysis. Pyruvate kinase converts phosphoenolpyruvate (PEP) to pyruvate and is an important regulatory step of glycolysis that is irreversible under most physiological conditions. When this enzyme was isolated from dehydrated *X. laevis* it showed increased affinity for PEP, which was correlated with decreased phosphorylation status compared to control animals (Dawson et al. 2018). The authors showed that the increased affinity for PEP may be a result of dephosphorylation. Likewise, liver LDH was also examined for altered kinetics and phosphorylation status in response to dehydration (Katzenback et al. 2014). In this case, LDH was phosphorylated and the K_m for all substrates increased during dehydration, reducing the catalytic efficiency of this enzyme. This, however, was reversed in the presence of physiological levels of urea (150 mM) in LDH from dehydrated animals but not control animals. The authors suggested that phosphorylation during dehydration may be altering the kinetics of LDH as a preparatory mechanism for rehydration. When rehydration occurs, urea is

rapidly excreted leaving the animal with phosphorylated LDH that now has less affinity for pyruvate, allowing it to flow into the TCA cycle. These studies show active modification of glycolytic enzymes that supports increased anaerobic glycolytic efficiency in the liver, which is necessary for processes such as the urea cycle.

In contrast to liver, skeletal muscle appears to be more dormant during dehydration. While dehydrating *X. laevis* assume a water conserving pose and remain stationary. This presumably results in lower metabolic requirements relative to the liver. Hexokinase, the gateway enzyme of glycolysis regulates the entry of glucose into this pathway and locks glucose in the form of glucose-6-phosphate in skeletal muscle cells. When isolated from dehydrated *X. laevis*, hexokinase showed a lower affinity for glucose but a higher V_{\max} (Childers and Storey 2016). Hexokinase also had reduced threonine phosphorylation and increased serine phosphorylation that was modelled to show an affect on glucose binding, potentially explaining the increased K_m during dehydration. In this case, hexokinase may be regulating glucose entry into glycolysis, especially since plasma glucose levels are elevated in response to dehydration (Malik and Storey 2009a). Similar to the liver, pyruvate kinase and LDH have also been analyzed in skeletal muscle. In this tissue, pyruvate kinase showed no altered kinetics (Dawson et al. 2018), whereas similar to the liver, LDH had increased affinity for pyruvate and elevated V_{\max} during dehydration (Childers and Storey 2019). Interestingly, phosphorylation of LDH decreased in skeletal muscle, the opposite of LDH in the liver, and therefore if this post-translational modification is responsible for the altered kinetics, it may be a different phosphorylation site to the one seen in liver LDH. Regardless, the results from glycolytic enzymes in

skeletal muscle show active regulation of anaerobic glycolysis which may conserve energy and reduce oxidation of plasma glucose unnecessarily.

1.2.3 Transcriptional and translational response to dehydration

Studies involving transcription factors and signaling cascades were the first to examine differential gene expression in the context of *X. laevis* dehydration. The members of the ERK signaling cascade, which have significant control over the cell cycle and proliferation, are significantly upregulated during dehydration, including increased levels of active phospho-ERK2 in skeletal muscle, liver, lung, heart, and kidney (Malik and Storey 2009a). These changes correlate with increased phosphorylation of the ERK2 targets p90 ribosomal S6 kinase (RSK) and transcription factor STAT3, and further correlation with the RSK target S6 ribosomal protein. While this study did not examine downstream targets of the active transcription factors, the increased protein and phosphorylation levels suggest modulation of central transcription-controlling signaling cascades in response to dehydration.

Another study examined the NRF2 transcription factor and found a tissue specific increase in transcript abundance, protein levels, and nuclear localization in response to dehydration (Malik and Storey 2009b). This transcription factor is known to induce the expression of many antioxidant enzymes and protect against oxidative injury. The glutathione S-transferase (GST) family of antioxidant enzymes is a known downstream target of NRF2 (Chanas et al. 2002), and indeed a correlation between elevated NRF2 and increased members of the GST family was reported in this study, suggesting transcriptional control over antioxidant activity during dehydration. These results were bolstered by a related study examining the FOXO1 transcription factor and downstream

antioxidant enzymes MnSOD and catalase (Malik and Storey 2011). In the liver from dehydrated *X. laevis*, the inactive phosphorylated form of FOXO1 was reduced and FOXO1 increased localization to the nucleus and showed a 57% higher DNA binding activity than in controls. Furthermore, both MnSOD and catalase showed increased transcript and protein levels in the liver as predicted by elevated FOXO1 activity, further demonstrating the transcriptional and translational response to dehydration in *X. laevis*.

As indicated above, pyruvate kinase from dehydrated *X. laevis* showed multiple modifications to its properties under dehydrating conditions including alterations to its kinetic parameters (a higher affinity for phosphoenolpyruvate and increased activation by fructose-1,6-bisphosphate) as well as transcriptional regulation (higher abundance of *pklr* transcripts as measured by RT-PCR) (Dawson et al. 2018). The authors suggested these adaptations facilitate anaerobic glycolysis, but more importantly is that even enzymes under kinetic modification are also modulated at the transcription level in a dehydration responsive manner.

1.3 The need for internal reference controls

As discussed, *X. laevis* is adapted to survive severe dehydration in part by changes in gene expression. The gold standard for measuring mRNA transcripts is reverse-transcription quantitative polymerase chain reaction (RT-qPCR) for its sensitivity and accuracy across multiple orders of magnitude of mRNA copy numbers. As the name implies, the mRNA in each experimental sample is reverse transcribed into cDNA prior to qPCR, which while necessary, can introduce error at the multiple stages of sample handling, preparation, and quantification.

For proper quantification of relative mRNA levels between two experimental conditions, an internal reference control is needed to account for variability of cDNA levels that are not based on the variability of the underlying biology (in this case actual mRNA levels). With methods such as Western blotting, the current standard for an internal reference control is simply a measure of total protein loaded into the well, commonly by Coomassie staining. This has replaced the use of “housekeeping” genes due to improved accuracy and lower cost (Welinder and Ekblad 2011). Unfortunately, the simplicity of this approach does not translate to RT-qPCR since an accurate measure of total mRNA levels in a RT-qPCR reaction is not feasible. In this case, the standard approach is to use one or more genes that are known to be stably expressed under the experimental conditions as a control for the errors introduced. This is done because the reference gene undergoes all the same manipulations that the genes of interest undergo, therefore, any introduced variability that would affect quantification can be accounted for.

Like Western blotting in the past, a “housekeeping” gene like *gapdh* or a cytoskeletal gene such as α -tubulin are used for their presumed expression stability under a variety of experimental manipulations. This presents a circular problem where if the quantification of gene A is used to normalize the quantification of gene B, how do you validate the quantification of gene A without another validated internal control? Over time it has been shown that the assumption that certain genes are stably expressed is incorrect in a wide range of experimental conditions (Kozera and Rapacz 2013). This is probably the case with studying dehydration tolerance in *X. laevis*. We know that glycolysis is heavily regulated during dehydration, including differential expression of glycolytic enzymes

(Dawson et al. 2018), suggesting that *gapdh* may be an inappropriate choice. Likewise, dehydration may lead to altered expression of cytoskeletal genes like α -tubulin since water is drawn out of the cells, potentially leading to changes in cell volume and structural strains.

A method other than the assumption of expression stability is therefore needed to validate potential reference genes for a given experimental setup. Once validated these genes can be used to normalize target values drawn from RT-qPCR analysis. If applied to the study of dehydration tolerance of *X. laevis*, the validation of reference genes will give more accurate measures of gene transcript levels, and therefore a better assessment of potential mechanisms contributing to dehydration tolerance.

1.4 Potential regulators of dehydration tolerance

Most cellular functions are carried out by the activity of proteins. The activity level of a given protein in a cell is a function of all the regulatory mechanisms that control the expression and function of that protein, spanning from the information encoded in the genome, through its transcription, translation, and post-translational modification, among others. The function of proteins contributes to the dehydration tolerance of *X. laevis*, therefore, each regulatory mechanism controlling the expression and function of proteins has the possibility of contributing to dehydration tolerance. This is evident in the literature, as discussed above, where transcriptional regulation (Malik and Storey 2009a, b; Dawson et al. 2018), post-transcriptional regulation (Wu et al. 2013; Luu and Storey 2015), and post-translational regulation (Katzenback et al. 2014; Childers and Storey 2016, 2019; Dawson et al. 2018) appear to be active in response to dehydration in *X. laevis*.

While some regulatory mechanisms have been examined previously, many have not been explored within the context of dehydration tolerance in *X. laevis*. This thesis aims to examine some of these unexplored or underexplored mechanisms (**Figure 1.2**).

Epigenetic mechanisms such as DNA methylation and histone modifications have previously been shown to change in animals adapted to diverse environmental stressors, including anurans (Hawkins and Storey 2018). The information in the genome gives instructions as to *what* is expressed, whereas epigenetics mechanisms comprise an additional layer of regulatory information on top of the genome which dictates *when* genes are expressed. While epigenetics controls the transcription of genes, microRNA control which transcribed genes are translated. These short RNA molecules can be cheaply expressed and negatively regulate mRNA translation in a targeted, precise manner. MicroRNA have shown to be differentially expressed in response to dehydration, although it has only been examined on a small scale (Wu et al. 2013; Luu and Storey 2015), whereas a more comprehensive characterization of the microRNAome may give insights into the types of pathways and processes involved. After a transcript is translated and the protein is formed, its activity can be modulated through post-translational modifications, phosphorylation being the most well characterized. Some studies have shown differential phosphorylation affecting pathways such as glycolysis (Katzenback et al. 2014; Childers and Storey 2016, 2019; Dawson et al. 2018), but like microRNA, protein phosphorylation has only been investigated on a protein-by-protein basis. A complete characterization of differential protein phosphorylation would help identify important players contributing to dehydration tolerance.

1.4.1 DNA methylation

DNA methylation refers to the addition of a methyl group to cytosine residues, typically at the 5th position carbon (5mC). Although 5mC is a chemically stable covalent modification, it is reversible and does not affect the interpretations of the underlying genetic code (Wang et al. 2015). S-adenosyl-methionine (SAM) is the universal methyl donor for DNA methylation, resulting in the formation of S-adenosyl-homocysteine (SAH) after transfer. The enzymes that catalyze the addition of the methyl groups are DNA methyltransferases (DNMTs) and fall into two categories, maintenance methyltransferases and *de novo* methyltransferases. DNMT1 has a high specificity for hemi-methylated DNA (Hermann et al. 2004), which occurs during DNA replication when the template strand contains DNA methylation patterns and the newly synthesized strand is unmodified and, thus, DNMT1 is the primary maintenance methyltransferase that copies extant patterns onto the new strand (Vertino et al. 2002; Schermelleh et al. 2007).

The *de novo* DNA methyltransferases DNMT3a and DNMT3b are the enzymes that catalyze the creation of new methylation patterns. Signals are propagated through standard cell signaling pathways and signaling molecules, transcription factors, and other chromatin modifying machinery to guide these enzymes to regions of DNA to be methylated. Besides requiring CpG sites, DNMT3a and DNMT3b appear to be sequence neutral (Dodge et al. 2002), and thus are flexible to methylate a wide variety of sequence contexts depending on their interacting partners. A third member of the DNMT3 group is DNMT3L which does not have catalytic activity, but acts as a regulatory factor and forms complexes with DNMT3a, DNMT3b and various transcription factors (Hata et al. 2002; Chedin et al. 2002; Vlachogiannis et al. 2015).

Prior to discovery of demethylating enzymes, it was thought that demethylation only occurred through dilution of 5mC bases by inhibition of maintenance methylation throughout replication cycles (Ooi and Bestor 2008). While passive dilution is one demethylation pathway, it was found that there are non-negligible levels of oxidized forms of 5mC which opened the door for more active demethylation mechanisms. The ten-eleven translocation (TET) family of enzymes serially oxidize 5mC to form 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Ito et al. 2011). Subsequently, 5fC and 5caC can then be excised by thymine DNA glycosylase (TDG) and the base excision repair pathway returns it to an unmodified cytosine (He et al. 2011; Maiti and Drohat 2011). Alternatively, once in the 5hmC form, maintenance methylation no longer recognizes this modification, and if it is not actively excised it will be passively diluted in actively replicating cell types (Wu and Zhang 2017).

The function of DNA methylation relies on the fact that the methyl group of 5mC is positioned within the major groove of the DNA helix, making it accessible to transcriptional machinery or repressors (Cheng and Blumenthal 2008). It is thought that it can sterically hinder the binding of transcriptional activators (Watt and Molloy 1988), while also creating binding ligands for transcriptional repressors such as methyl-binding domain (MBD) containing proteins (Dhasarathy and Wade 2008). The MBD is highly conserved among vertebrates, and within the proteins that contain it. There are five well characterized MBD containing proteins, methyl-CpG binding protein 2 (MeCP2) and MBD1-4. These proteins bind specifically to methylated CpGs, and recruit additional transcriptionally repressive proteins and chromatin modifiers such as histone deacetylases

(HDACs) (Nan et al. 1998) that result in compaction of the DNA and further reduce access to transcriptional activators.

There has been some disagreement as to whether DNA methylation and transcriptional repression is a chicken or the egg problem. While many studies regarding DNMT activity, MBD binding, and 5mC distribution are correlative, there is sparse evidence that in specific scenarios methylation appears after transcriptional silencing (Suzuki and Bird 2008), refuting its role as a causal player. In recent years genome editing tools have become more powerful and allow us to test problems such as this one. Multiple groups have created endonuclease dead Cas9 (dCas9) conjugated DNMT systems capable of programmatically methylating specific CpG sites (McDonald et al. 2016; Stepper et al. 2017). In both cases, these groups were able to specifically methylate targeted genes and repress transcription, resulting in some of the best evidence for a causal role of DNA methylation in transcriptional silencing.

Given the application to transcriptional regulation, it should come as no surprise that DNA methylation has been characterized to the greatest extent in the fields of developmental and cancer biology. Both fields have demonstrated the power and involvement of DNA methylation in any given cellular process, and thus other processes involving transcriptional reorganization, such as adaptations to environmental stressors, are obvious targets to study.

1.4.2 Histone modifications

DNA is wrapped around an octamer of histone proteins forming a nucleosome. Nucleosomes are the base units of chromatin that DNA in a dynamically reversible fashion. Chromatin can be packaged in a condensed and normally repressive form known

as heterochromatin, or in a transcriptionally permissive relaxed euchromatic form. The different conformations can occur simultaneously at different genomic loci depending both on chromatin remodeling proteins and the histones themselves (Bannister and Kouzarides 2011). As with DNA, histones are covalently modified to alter their structure and function. The C- and N- terminal tails of each histone protein, and to some extent the histone cores, can be modified post-translationally with the addition of numerous chemical groups. The resulting effect of each modification depends on its type, placement along the histone tails, and surrounding modifications - the aggregate of which has been called the *histone code* (Jenuwein and Allis 2001).

Histone lysine residues can be acetylated through the transfer of acetyl group from acetyl-CoA by histone acetyltransferases (HATs) and deacetylated by histone deacetylases (HDACs). Acetylation serves multiple functions: 1) it removes the positive charge from the lysine residue, reducing the histone-DNA interaction (Lee et al. 1993), 2) it creates a binding ligand for proteins to interact with (Dhalluin et al. 1999), and 3) it excludes the addition of other modifications, for example histone H3 lysine 9 (H3K9) can be acetylated to activate transcription or methylated to silence transcription, but not both simultaneously. Together these functions generally mean that acetylation produces a positive or permissive effect on transcription.

Histones lysine residues can also be methylated however, unlike acetylation, the methyl group does not alter the electrostatic properties of the histone protein. This lack of intrinsic effect compared to acetylation means that function of methylation is dependant on other interacting proteins. Histone lysine methylation is catalyzed by histone methyltransferases (HMTs) with SAM as the sole methyl-donor (Luka et al. 2009).

HMTs transfer the methyl group from SAM and add it to the lysine ϵ -amino group, making it mutually exclusive with acetylation (Seet et al. 2006). Most histone lysine methyltransferases contain a Su(var)3-9, Enhancer-of-zeste and Trithorax domain (Dillon et al. 2005), which is the main methyltransferase domain. In contrast to acetylation, up to three methyl groups can be added to each lysine residue, each with a possibly distinct function depending on the proteins that interact with it. Studies on specific methylated lysines reveal associations with transcriptional and conformation states (Kouzarides 2002). While there are many lysines that can be methylated, the residues with strong correlations to specific transcriptional outcomes are H3K4 and H3K9. For example, H3K4 mono-methylation (H3K4me) is present on active gene enhancers (Local et al. 2018), while H3K9me3 is found in silenced genes often co-occurring with DNA methylation and a lack of histone acetylation (Fuks 2005).

Histones are also dynamically modified in response to stress and environmental conditions. Studies on mice exposed to different social stressors show both global (Xu et al. 2018) and gene specific (Krause et al. 2015) histone modification patterns. Numerous studies have shown that exposure to toxicants such as benzene (Li et al. 2018), fluoride (Daiwile et al. 2018), chemical warfare agents (Simons et al. 2017), phthalates (Singh and Li 2012; Sun et al. 2013), and heavy metals (Ma et al. 2015; Gadhia et al. 2015; Eid and Zawia 2016) affect histone modifications. Interestingly, many of these studies show changes to levels of H3K4me3 as a common result, however, this may say more about experimental selection of histone modifications than convergent biological relevance. Several natural models for studying environmental stresses have also been examined in the context of histone modifications. The freeze-tolerant wood frog showed decreased

permissive histone modifications (H3K4me1 and H3K27me1) during freezing, suggesting a contribution from histone methylation to global transcriptional repression (Hawkins and Storey 2018). The red-eared slider showed analogous results with an increase in the repressive H3K9me3, and H3K9 methyltransferase protein and activity levels (Wijenayake et al. 2018), and decreases in acetylated histones under anoxic conditions (Krivoruchko and Storey 2010). Hibernating ground squirrels show distinct changes to histone modifying enzymes and modification levels across torpor cycles (Morin and Storey 2006; Biggar and Storey 2014; Rouble and Storey 2015; Tessier et al. 2017), supporting previous findings that transcription is repressed during hibernation (van Breukelen and Martin 2002; Heldmaier et al. 2004). While the exact function of each histone modification with respect to a given stress is still being determined, it is likely that they are part of the larger gene regulatory response, particularly in the case of adaptation to environmental conditions honed over evolutionary time.

1.4.3 MicroRNA

MicroRNA are short non-coding RNA that silence translation of target mRNA. MicroRNA typically bind to the 3'-untranslated region (UTR) of mRNA and depending on the degree of complementarity, result in mRNA degradation (in the case of complete complementarity) or translational repression and storage of the mRNA transcript into cellular compartments such as P-bodies (He and Hannon 2004; Liu et al. 2005).

Like mRNA, microRNA are transcribed by RNA Pol II-associated transcriptional machinery (Lee et al. 2004). Most microRNA coding regions are within introns of coding and non-coding genes. This means that they can share the promoter of the outer gene, but often have multiple transcription start sites and therefore separate promoters (Ozsolak et

al. 2008). MicroRNA are initially transcribed as a primary microRNA (pri-microRNA) (Lee et al. 2002). These structures contain a stem-loop of dsRNA containing the final mature sequence, as well as ssRNA tails at either end. The nuclear ribonuclease Drosha and its cofactor DGCR8 cleave the pri-microRNA approximately 20-22 bp away from the loop of the hairpin to form what is known as a precursor microRNA (pre-microRNA) (Han et al. 2004). The pre-microRNA is then translocated out of the nucleus and into the cytoplasm by exportin 5 and RanGTP (Bohnsack et al. 2004). In the cytoplasm a second ribonuclease, Dicer, cleaves the terminal loop from the hairpin leaving a 20-22 bp dsRNA molecule (Ketting et al. 2001). This RNA duplex contains a “guide” strand, which is fully- or partially complementary to the target mRNA, and a “passenger” strand which is generally not biologically active. A mature RNA-induced silencing complex (RISC) is created with one of the members of the Argonaut (AGO) family, with the help of the HSC70-HSP90 chaperone, loads the pre-microRNA duplex and discards the passenger strand (Iwasaki et al. 2010). Which strand is the guide strand is determined by thermodynamic stability (selecting the strand with the less stable 5' end) and a terminal uracil preference by the AGO protein (Wu et al. 2009; Hu et al. 2009).

Dehydration responsive microRNA expression has been analyzed in small-scale, targeted studies (Wu et al. 2013; Luu and Storey 2015). The first microRNA study examined tissue-specific responses by a small group of *X. laevis* microRNAs with high homology to known human microRNA (Wu et al. 2013). Bioinformatic analysis of the differentially expressed microRNAs found two potentially targeted processes 1) MAPK signaling cascades, a known dehydration responsive pathway (Malik and Storey 2009a), and 2) expression of passive membrane solute carriers that would respond to the

increased cellular and plasma osmolarity. A second study examined microRNA specifically in *X. laevis* brain (Luu and Storey 2015) and this study found several downregulated microRNAs that indicated that neuroprotective pathways, such as axon guidance and long-term potentiation, were activated under dehydrated conditions.

The action of microRNA is a powerful post-transcriptional regulatory mechanism. They are metabolically “cheap” to produce, fast-acting, and are capable of transcriptome wide targeting. For these reasons it comes as no surprise that microRNA have been implicated in virtually every cellular and metabolic process. Adaptation to adverse environmental conditions are comprised of these processes and pathways, and thus microRNA regulation is a suitable candidate for studying the molecular mechanisms underlying dehydration tolerance.

1.4.4 Reversible protein phosphorylation

Reversible protein phosphorylation is a cornerstone of metabolic regulation in animals that survive extreme environmental conditions (Storey and Storey 1990; Cowan and Storey 2003; Hawkins and Storey 2020). Post-translational modification like phosphorylation can turn on and off entire metabolic pathways, activate or deactivate proteins, and direct the subcellular localization of numerous proteins. Knowing the phosphorylation status of proteins therefore can tell you much about what is going on in the cell.

Within the context of adaptations to environmental conditions, reversible phosphorylation of proteins can result in unique control over pathways and metabolism. A particularly clear example of this is the glucose metabolism shown by the freeze tolerant wood frog (*Rana sylvatica*). This animal catabolizes huge amounts of glycogen

to build up cryoprotectant glucose within minutes of exposure to freezing conditions. Under the regulation of canonical vertebrate glycogenolysis, the breakdown of glycogen would be allosterically inhibited by the rise in glucose-6-phosphate levels and some of the resulting glucose would flux through glycolysis. In this case, when the wood frog begins to freeze a new glycogenolytic regulatory program runs where glycogen is continuously broken down into glucose, and that glucose is metabolically blocked from entering glycolysis, potentially due to novel phosphorylation sites across multiple enzymes including glycogen phosphorylase and phosphofructokinase (Hawkins et al. 2019).

In *X. laevis*, the effect of reversible phosphorylation has mostly been studied with regard to glucose metabolism genes, namely hexokinase, pyruvate kinase, and LDH (Katzenback et al. 2014; Childers and Storey 2016, 2019; Dawson et al. 2018). The differential phosphorylation of these enzymes has been correlated with changes in their kinetics that appear to promote or facilitate anaerobic metabolism during dehydration. As discussed above, MAPK proteins have also shown to be differentially phosphorylated leading to the transcriptional regulation of downstream genes (Malik and Storey 2009a). In each of these studies, specifically targeted proteins were examined which necessarily limits the scope of conclusions that can be drawn. By taking a comprehensive approach to the study of phosphorylation changes under a given experimental condition an incredible amount of information can be gleaned about cellular processes and pathways (Sacco et al. 2018). The application of phosphoproteomic approaches to the study of dehydration tolerance would therefore elucidate potentially unknown mechanisms that would otherwise remain hidden from the previously taken targeted approaches.

1.5 Objectives

This thesis aims to increase our understanding of dehydration tolerance in *X. laevis* by first establishing important controls in the form of stably expressed reference genes, then using those genes as tools in the examination of gene expression regulatory mechanisms that may contribute to dehydration tolerance as outlined in the following objectives:

1.5.1 Objective 1 - Determine suitable reference genes for RT-qPCR analysis of dehydration tolerance in multiple tissues of *Xenopus laevis*

The use of RT-qPCR is ubiquitous in molecular biology studies. Its power to quantify mRNA levels spanning many orders of magnitude combined with the many opportunities to introduce error either in sample handling, preparation, or measurement, means use of a proper internal control is crucial to accurate quantification. Internal controls to account for variations in cDNA levels not underpinned by the biology of a system should be determined for each experimental setup. While internal controls in the form of reference genes have been determined for developmental studies of *X. laevis* (Mughal et al. 2018), the same has not been done for studies of dehydration tolerance. This objective, addressed in **Chapter 2** of this thesis, is therefore to determine proper internal controls by validating potential reference genes. The reference genes determined in this chapter are then used in subsequent chapters where RT-qPCR is employed in the study of dehydration tolerance in *X. laevis* making this objective a prerequisite for future analysis.

1.5.2 Objective 2 - Quantify global changes of DNA methylation and histone modifications in *Xenopus laevis* during dehydration

The dehydration tolerance of *X. laevis* is predicated in part on changes to gene expression in a tissue specific manner. Many mechanisms exist for selectively regulating which genes are expressed in response to a given signal from the environment. Potential mechanisms include epigenetic mechanism such as DNA methylation and histone modifications, both of which have been shown to be involved in the adaptational response to environmental stresses in numerous other animals. This objective, explored in **Chapter 3**, is to quantify global changes in epigenetic marks such as 5mC, 5hmC, and various histone modifications, along with the expression levels of the enzymes responsible for their addition and removal in *X. laevis* during dehydration. The results of these experiments can elucidate if and which epigenetic mechanisms may be contributing to dehydration tolerance and whether there may be tissue-specific effects on gene expression mediated by these mechanisms.

1.5.3 Objective 3 - Profile changes to the cardiac microRNAome in response to dehydration in *Xenopus laevis*

MicroRNA are energetically cheap to produce and can quickly and specifically adjust the translation of mRNA transcripts. These attributes make microRNA perfectly suited for the cell to use in response to changing environmental conditions where a large-scale reprogramming of gene expression is necessary. Such is the case during dehydration in *X. laevis*, and therefore this objective and **Chapter 4** is to comprehensively profile the microRNAome in the heart of this animal. The heart was chosen for this study as it is relatively unexplored from a molecular biology approach in the context of dehydration

tolerance, even though it physiologically active and important for survival under these conditions. Determining the differentially expressed microRNA, and the genes and pathways they target will give indications of potentially unknown adaptational responses that allow the heart to sustain high activity and contribute to dehydration tolerance in *X. laevis*.

1.5.4 Objective 4 - Analyze the liver and skeletal muscle phosphoproteome of *Xenopus laevis* during dehydration

The phosphorylation status of a protein can give several indications as to its activity, function, localization, interacting partners, and more. In instances where the result of phosphorylation or dephosphorylation of a specific residue is known, measuring changes in phosphorylation status between experimental conditions can elucidate specific cellular processes and pathways implicated. For example, measuring greater levels of phosphorylated glycogen phosphorylase indicates glycogenolysis is likely occurring under a given experimental condition, since phosphorylation of glycogen phosphorylase activates this enzyme. On the contrary, even when no information about the function of a specific phospho-residue is available, differences in phosphorylation of a protein implicate it in the scenario since phosphorylation will modulate one or more parameters of that protein. These points contribute to the utility of phosphoproteomic approaches which detect proteome wide phosphorylation changes of known and unknown function. Previous targeted studies of protein phosphorylation have been fruitful in revealing dehydration responsive adaptations, therefore this objective as part of **Chapter 5** uses phosphoproteomic analysis to reveal potential dehydration tolerance adaptation in *X. laevis* on a large scale. By comparing the phosphoproteome and differentially

phosphorylated proteins in the liver and skeletal muscle, I can demonstrate tissue level differences in this animal's response to dehydration, specific instances in each tissue, and broaden our understanding of the adaptations involved.

1.6 References

- Balinsky JB, Choritz EL, Coe CG, van der Schans GS (1967) Amino acid metabolism and urea synthesis in naturally aestivating *Xenopus laevis*. *Comp Biochem Physiol* 22:59–68 . doi: 10.1016/0010-406x(67)90166-1
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21:381–95 . doi: 10.1038/cr.2011.22
- Biggar Y, Storey KB (2014) Global DNA modifications suppress transcription in brown adipose tissue during hibernation. *Cryobiology* 69:333–8 . doi: 10.1016/j.cryobiol.2014.08.008
- Bohnsack MT, Czaplinski K, Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 10:185–191 . doi: 10.1261/rna.5167604
- Boothby TC, Tapia H, Brozena AH, et al (2017) Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. *Mol Cell* 65:975-984.e5 . doi: 10.1016/j.molcel.2017.02.018
- Chanas SA, Jiang Q, McMahon M, et al (2002) Loss of the Nrf2 transcription factor causes a marked reduction in constitutive and inducible expression of the glutathione S-transferase *Gsta1*, *Gsta2*, *Gstm1*, *Gstm2*, *Gstm3* and *Gstm4* genes in the livers of male and female mice. *Biochem J* 365:405–16 . doi: 10.1042/BJ20020320
- Chedin F, Lieber MR, Hsieh C-L (2002) The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A* 99:16916–21 . doi: 10.1073/pnas.262443999
- Cheng X, Blumenthal RM (2008) Mammalian DNA Methyltransferases: A Structural Perspective. *Structure* 16:341–350 . doi: 10.1016/j.str.2008.01.004
- Childers CL, Storey KB (2016) Post-translational regulation of hexokinase function and protein stability in the aestivating frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Childers CL, Storey KB (2019) Purification and characterization of a urea sensitive lactate dehydrogenase from skeletal muscle of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B Biochem Syst Environ Physiol* 189:271–281 . doi: 10.1007/s00360-018-1200-3
- Cowan KJ, Storey KB (2003) Mitogen-activated protein kinases: new signaling pathways functioning in cellular responses to environmental stress. *J Exp Biol* 206:1107–15 . doi: 10.1242/jeb.00220
- Daiwile AP, Sivanesan S, Tarale P, et al (2018) Role of fluoride induced histone trimethylation in development of skeletal fluorosis. *Environ Toxicol Pharmacol* 57:159–165 . doi: 10.1016/j.etap.2017.12.015

- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- Dhalluin C, Carlson JE, Zeng L, et al (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399:491–496 . doi: 10.1038/20974
- Dhasarathy A, Wade PA (2008) The MBD protein family-Reading an epigenetic mark? *Mutat Res - Fundam Mol Mech Mutagen* 647:39–43 . doi: 10.1016/j.mrfmmm.2008.07.007
- Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 6:227 . doi: 10.1186/gb-2005-6-8-227
- Dodge JE, Ramsahoye BH, Wo ZG, et al (2002) De novo methylation of MMLV provirus in embryonic stem cells: CpG versus non-CpG methylation. *Gene* 289:41–8 . doi: 10.1016/S0378-1119(02)00469-9
- Eid A, Zawia N (2016) Consequences of lead exposure, and it's emerging role as an epigenetic modifier in the aging brain. *Neurotoxicology* 56:254–261 . doi: 10.1016/j.neuro.2016.04.006
- Erkut C, Penkov S, Khesbak H, et al (2011) Trehalose renders the dauer larva of *Caenorhabditis elegans* resistant to extreme desiccation. *Curr Biol* 21:1331–6 . doi: 10.1016/j.cub.2011.06.064
- Fuks F (2005) DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* 15:490–5 . doi: 10.1016/j.gde.2005.08.002
- Gadhia SR, O'Brien D, Barile FA (2015) Cadmium affects mitotically inherited histone modification pathways in mouse embryonic stem cells. *Toxicol In Vitro* 30:583–92 . doi: 10.1016/j.tiv.2015.11.001
- Han J, Lee Y, Yeom K, et al (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18:3016–27 . doi: 10.1101/gad.1262504
- Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129:1983–93
- Hawkins LJ, Storey KB (2018) Histone methylation in the freeze-tolerant wood frog (*Rana sylvatica*). *J Comp Physiol B* 188:113–125 . doi: 10.1007/s00360-017-1112-7
- Hawkins LJ, Storey KB (2020) Advances and applications of environmental stress adaptation research. *Comp Biochem Physiol A Mol Integr Physiol* 240:110623 . doi: 10.1016/j.cbpa.2019.110623
- Hawkins LJ, Wang M, Zhang B, et al (2019) Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposures in a freeze tolerant frog. *Comp Biochem Physiol Part D Genomics Proteomics* 30:1–13 . doi: 10.1016/j.cbd.2019.01.009

- He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5:522–31 . doi: 10.1038/nrg1379
- He Y-F, Li B-Z, Li Z, et al (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333:1303–7 . doi: 10.1126/science.1210944
- Heldmaier G, Ortman S, Elvert R (2004) Natural hypometabolism during hibernation and daily torpor in mammals. *Respir Physiol Neurobiol* 141:317–29 . doi: 10.1016/j.resp.2004.03.014
- Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* 279:48350–9 . doi: 10.1074/jbc.M403427200
- Hillman SS (1978a) The roles of oxygen delivery and electrolyte levels in the dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Hillman SS (1978b) Some effects of dehydration on internal distributions of water and solutes in *Xenopus laevis*. *Comp Biochem Physiol Part A Physiol* 61:303–307
- Hillman SS, Sommerfeldt RW (1981) Microsphere studies of amphibian systemic blood flow redistribution during dehydration, hypovolemia, and salt load. *J Exp Zool* 218:305–308
- Hu HY, Yan Z, Xu Y, et al (2009) Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* 10:413 . doi: 10.1186/1471-2164-10-413
- Ito S, Shen L, Dai Q, et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333:1300–3 . doi: 10.1126/science.1210597
- Iwasaki S, Kobayashi M, Yoda M, et al (2010) Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Mol Cell* 39:292–299 . doi: 10.1016/j.molcel.2010.05.015
- Jenuwein T, Allis CD (2001) Translating the histone code. *Science* 293:1074–80 . doi: 10.1126/science.1063127
- Jokumsen A, Weber RE (1980) Haemoglobin-oxygen binding properties in the blood of *Xenopus laevis*, with special reference to the influences of aestivation and of temperature and salinity acclimation. *J Exp Biol* 86:19–37
- Karimi K, Fortriede JD, Lotay VS, et al (2018) Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res* 46:D861–D868 . doi: 10.1093/nar/gkx936
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Ketting RF, Fischer SEJ, Bernstein E, et al (2001) Dicer functions in RNA interference

- and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15:2654–9 . doi: 10.1101/gad.927801
- Kouzarides T (2002) Histone methylation in transcriptional control. *Curr Opin Genet Dev* 12:198–209
- Kozera B, Rapacz M (2013) Reference genes in real-time PCR. *J Appl Genet* 54:391–406 . doi: 10.1007/s13353-013-0173-x
- Krause L, Haubold B, Börsch-Haubold AG (2015) Social exclusion changes histone modifications H3K4me3 and H3K27ac in liver tissue of wild house mice. *PLoS One* 10:e0133988 . doi: 10.1371/journal.pone.0133988
- Krivoruchko A, Storey KB (2010) Epigenetics in anoxia tolerance: a role for histone deacetylases. *Mol Cell Biochem* 342:151–161 . doi: 10.1007/s11010-010-0479-5
- Lee DY, Hayes JJ, Pruss D, Wolffe AP (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72:73–84 . doi: 10.1016/0092-8674(93)90051-Q
- Lee Y, Jeon K, Lee J-T, et al (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* 21:4663–70 . doi: 10.1093/emboj/cdf476
- Lee Y, Kim M, Han J, et al (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23:4051–4060 . doi: 10.1038/sj.emboj.7600385
- Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245 . doi: 10.1093/nar/gkw290
- Li J, Xing X, Zhang X, et al (2018) Enhanced H3K4me3 modifications are involved in the transactivation of DNA damage responsive genes in workers exposed to low-level benzene. *Environ Pollut* 234:127–135 . doi: 10.1016/j.envpol.2017.11.042
- Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7:719–723 . doi: 10.1038/ncb1274
- Local A, Huang H, Albuquerque CP, et al (2018) Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat Genet* 50:73–82 . doi: 10.1038/s41588-017-0015-6
- Loveridge PJ (1976) Strategies of water conservation in Southern African frogs. *Zool Africana* 11:319–333
- Luka Z, Mudd SH, Wagner C (2009) Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. *J Biol Chem* 284:22507–22511 . doi: 10.1074/jbc.R109.019273
- Luu BE, Storey KB (2015) Dehydration triggers differential microRNA expression in *Xenopus laevis* brain. *Gene* 573:64–9
- Ma L, Bai Y, Pu H, et al (2015) Histone Methylation in Nickel-Smelting Industrial Workers. *PLoS One* 10:e0140339 . doi: 10.1371/journal.pone.0140339
- Maiti A, Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-

- formylcytosine and 5-carboxylecytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* 286:35334–8 . doi: 10.1074/jbc.C111.284620
- Malik AI, Storey KB (2009a) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Malik AI, Storey KB (2009b) Activation of antioxidant defense during dehydration stress in the African clawed frog. *Gene* 442:99–107 . doi: 10.1016/j.gene.2009.04.007
- Malik AI, Storey KB (2011) Transcriptional regulation of antioxidant enzymes by FoxO1 under dehydration stress. *Gene* 485:114–9 . doi: 10.1016/j.gene.2011.06.014
- McDonald JI, Celik H, Rois LE, et al (2016) Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. *Biol Open* 5:866–874 . doi: 10.1242/bio.019067
- Moody SA, Sater AK, Alfandari D, et al (2016) *Xenopus* Community White Paper 2016
- Morin P, Storey KB (2006) Evidence for a reduced transcriptional state during hibernation in ground squirrels. *Cryobiology* 53:310–8 . doi: 10.1016/j.cryobiol.2006.08.002
- Mughal BB, Leemans M, Spirhanzlova P, et al (2018) Reference gene identification and validation for quantitative real-time PCR studies in developing *Xenopus laevis*. *Sci Rep* 8:496 . doi: 10.1038/s41598-017-18684-1
- Nan X, Ng HH, Johnson CA, et al (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393:386–389 . doi: 10.1038/30764
- Ooi SKT, Bestor TH (2008) The colorful history of active DNA demethylation. *Cell* 133:1145–8 . doi: 10.1016/j.cell.2008.06.009
- Ozsolak F, Poling L, Wang Z, et al (2008) Chromatin structure analyses identify miRNA promoters. *Genes Dev* 23:3172–3183 . doi: 10.1101/gad.1706508.sion
- Rouble AN, Storey KB (2015) Characterization of the SIRT family of NAD⁺-dependent protein deacetylases in the context of a mammalian model of hibernation, the thirteen-lined ground squirrel. *Cryobiology* 71:334–43 . doi: 10.1016/j.cryobiol.2015.08.009
- Sacco F, Perfetto L, Cesareni G (2018) Combining Phosphoproteomics Datasets and Literature Information to Reveal the Functional Connections in a Cell Phosphorylation Network. *Proteomics* 18:e1700311 . doi: 10.1002/pmic.201700311
- Schermelleh L, Haemmer A, Spada F, et al (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* 35:4301–12 . doi: 10.1093/nar/gkm432
- Seet BT, Dikic I, Zhou MM, Pawson T (2006) Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol* 7:473–483 . doi: 10.1038/nrm1960
- Session AM, Uno Y, Kwon T, et al (2016) Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature* 538:336–343 . doi: 10.1038/nature19840

- Simons T, Steinritz D, Bölck B, et al (2017) Sulfur mustard-induced epigenetic modifications over time - a pilot study. *Toxicol Lett* 0–1 . doi: 10.1016/j.toxlet.2017.11.010
- Singh S, Li SSL (2012) Epigenetic effects of environmental chemicals bisphenol A and phthalates. *Int J Mol Sci* 13:10143–10153 . doi: 10.3390/ijms130810143
- Stepper P, Kungulovski G, Jurkowska RZ, et al (2017) Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res* 45:1703–1713 . doi: 10.1093/nar/gkw1112
- Storey KB (2015) Regulation of hypometabolism: insights into epigenetic controls. *J Exp Biol* 218:150–159
- Storey KB, Storey JM (2004) Metabolic rate depression in animals: Transcriptional and translational controls. *Biol Rev Camb Philos Soc* 79:207–233
- Storey KB, Storey JM (1990) Metabolic rate depression and biochemical adaptation in anaerobiosis, hibernation and estivation. *Q Rev Biol* 65:145–174
- Sun Y, Guo Z, Iku S, et al (2013) Diethyl phthalate enhances expression of SIRT1 and DNMT3a during apoptosis in PC12 cells. *J Appl Toxicol* 33:1484–92 . doi: 10.1002/jat.2816
- Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9:465–76 . doi: 10.1038/nrg2341
- Tessier SN, Luu BE, Smith JC, Storey KB (2017) The role of global histone post-translational modifications during mammalian hibernation. *Cryobiology* 75:28–36 . doi: 10.1016/j.cryobiol.2017.02.008
- van Breukelen F, Martin SL (2002) Reversible depression of transcription during hibernation. *J Comp Physiol B* 172:355–61 . doi: 10.1007/s00360-002-0256-1
- Vertino PM, Sekowski JA, Coll JM, et al (2002) DNMT1 is a component of a multiprotein DNA replication complex. *Cell Cycle* 1:416–23 . doi: 10.4161/cc.1.6.270
- Vlachogiannis G, Niederhuth CE, Tuna S, et al (2015) The Dnmt3L ADD Domain Controls Cytosine Methylation Establishment during Spermatogenesis. *Cell Rep* 10:944–956 . doi: 10.1016/j.celrep.2015.01.021
- Wang L, Zhou Y, Xu L, et al (2015) Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature* 523:621–5 . doi: 10.1038/nature14482
- Watt F, Molloy PL (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* 2:1136–1143 . doi: 10.1101/gad.2.9.1136
- Welinder C, Ekblad L (2011) Coomassie staining as loading control in Western blot analysis. *J Proteome Res* 10:1416–9 . doi: 10.1021/pr1011476
- Wijenayake S, Hawkins LJ, Storey KB (2018) Dynamic regulation of six histone H3 lysine (K) methyltransferases in response to prolonged anoxia exposure in a

- freshwater turtle. *Gene* 649:50–57 . doi: 10.1016/j.gene.2018.01.086
- Worland MR, Block W (1986) Survival and water loss in some Antarctic arthropods. *J Insect Physiol* 32:579–584 . doi: 10.1016/0022-1910(86)90075-2
- Wu CW, Biggar KK, Storey KB (2013) Dehydration mediated microRNA response in the African clawed frog *Xenopus laevis*. *Gene* 529:269–75
- Wu H, Ye C, Ramirez D, Manjunath N (2009) Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS One* 4: . doi: 10.1371/journal.pone.0007566
- Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 18:517–534 . doi: 10.1038/nrg.2017.33
- Xu H, Wang J, Zhang K, et al (2018) Effects of adolescent social stress and antidepressant treatment on cognitive inflexibility and Bdnf epigenetic modifications in the mPFC of adult mice. *Psychoneuroendocrinology* 88:92–101 . doi: 10.1016/j.psyneuen.2017.11.013

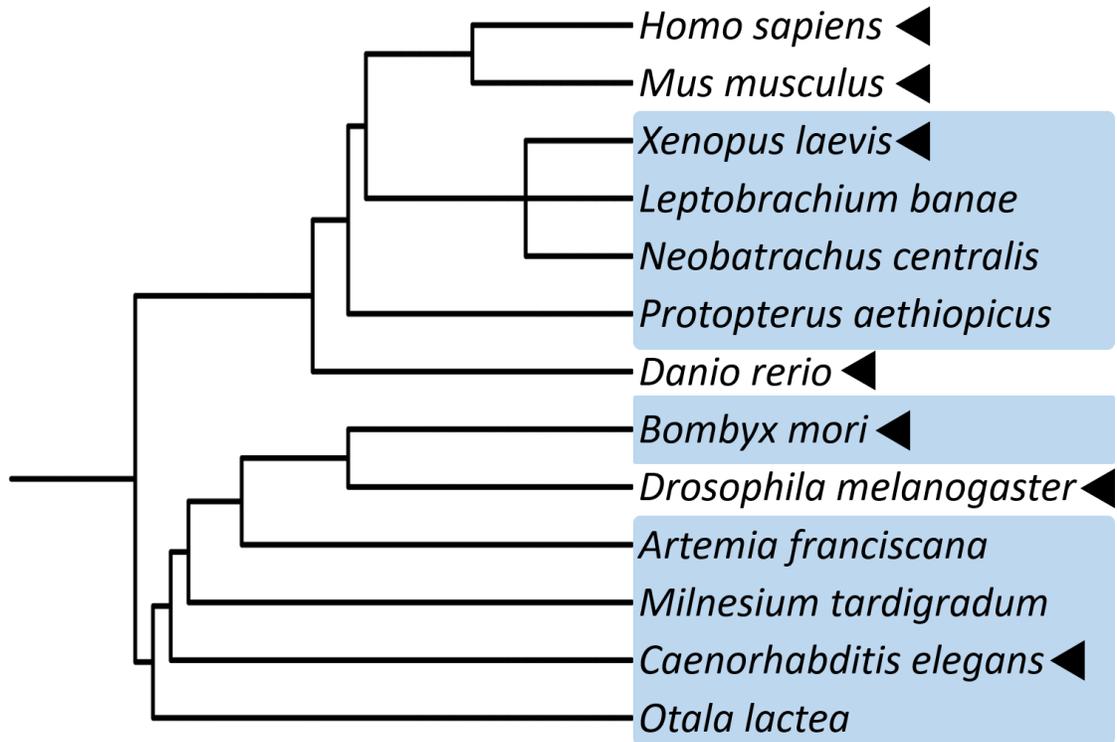


Figure 1.1 Phylogenetic tree of select model and dehydration tolerant animals. Species in blue are dehydration tolerant. Species with a black arrow have well-annotated genomes. *Xenopus laevis* is the most closely related species to humans that is dehydration tolerant (blue) and has a fully annotated genome (arrow). The phylogenetic tree was generated using Interactive Tree of Life (iTOL) (Letunic and Bork 2016).

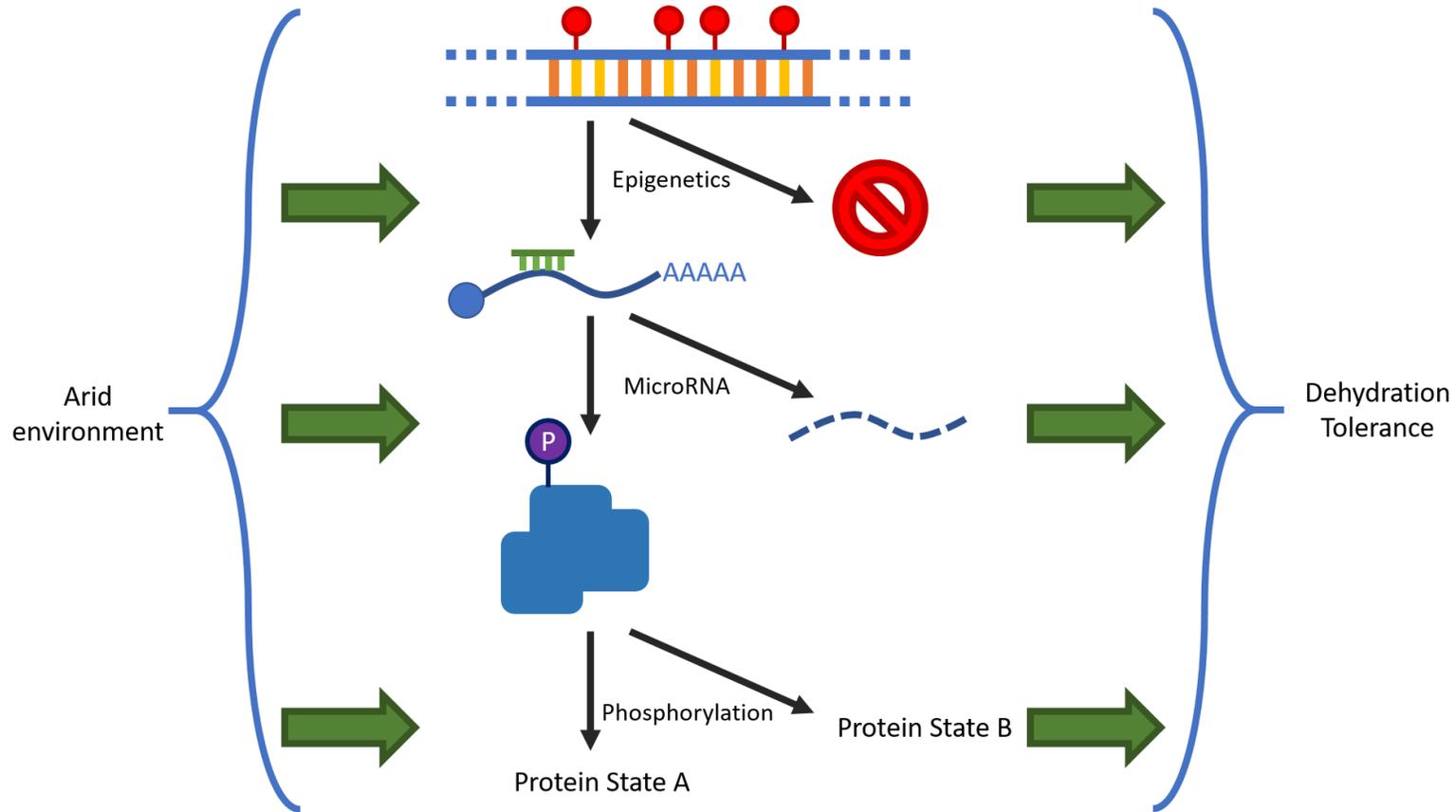


Figure 1.2 Select mechanisms of regulation potentially contributing to dehydration tolerance in *Xenopus laevis*. As the environment becomes arid and dehydration progresses, signals are transduced to a variety of regulatory mechanisms that enact adaptational responses enabling survival until rehydration occurs. The mechanisms of regulation studied here start at the epigenetic level where DNA methylation and histone modifications are examined, then an analysis of the microRNAome is undertaken, followed by phosphoproteomic analysis.

Chapter 2

Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*

Liam J. Hawkins¹, Bryan E. Luu¹, and Kenneth B. Storey¹

¹ Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

This study has been published in Gene Reports:

Hawkins, L. J., Luu, B. E., & Storey, K. B. (2018). Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. *Gene Reports*, 13, 192–198. <https://doi.org/10.1016/j.genrep.2018.10.006>

2.1 Abstract

Reverse-transcription quantitative polymerase chain reaction (RT-qPCR) is a common technique for quantifying gene expression. The reliability and reproducibility of RT-qPCR relies heavily on normalization of target values to an internal control to account for experimental variability. *Xenopus laevis*, is well-known as a model in embryology and developmental biology but, with a sequenced genome, is now attracting attention as a natural model for studying environmental stress tolerance in adult amphibians. Specifically, recent studies have focused on gene responses to dehydration, however, a systematic screening of reference genes has yet to be done. Here, we evaluate the stability of eight candidate reference genes (*actb.L*, *atp5bf1.L*, *canx.L*, *gapdh.S*, *rpl27.L*, *sdha.L*, *tbp.L*, *tbp.S*) in eight tissues (brain, dorsal skin, heart, kidney, liver, lung, skeletal muscle, and ventral skin) using three popular algorithms (geNorm, Normfinder, and BestKeeper) in dehydrated *X. laevis*. A tissue-specific composite rank of each gene was determined from the stability rankings of each algorithm, and the optimal number of reference genes was determined by the pairwise-variation of normalization factors using the top n and $n + 1$ genes. The following reference gene sets were determined for each tissue: *canx.L* and *rpl27.L* in brain, *rpl27.L* and *sdha.L* in dorsal skin, *actb.L* and *rpl27.L* in heart, *actb.L*, *gapdh.S*, *tbp.L*, *atp5f1b.L*, *tbp.S* and *rpl27.L* in kidney, and *atp5f1b.L* in liver, *actb.L* and *rpl27.L* in lung, *gapdh.S* and *rpl27.L* in skeletal muscle, and *rpl27.L* and *canx.L* in ventral skin. The present study recommends that these genes be used as internal controls in future studies on dehydration tolerance in *X. laevis*.

2.2 Introduction

Xenopus laevis has been used as a model organism in cellular and developmental biology studies for more than 50 years. There are many characteristics that make it an attractive model for research, such as ease of handling, housing, and husbandry, large brood and egg sizes, and accessible embryos. In fact, many major breakthroughs in the biological sciences were achieved using *X. laevis*, including Nobel Prize-winning contributions by John B. Gurdon, which demonstrated that somatic cells can be reprogrammed to become pluripotent (Gurdon 1962).

Whereas most work in *X. laevis* involves its use as a classic model organism, its interesting life history traits have been largely ignored. *X. laevis* is a primarily aquatic anuran that is native to southern Africa and inhabits ponds, lakes, and wetlands. These habitats are highly susceptible to seasonal drying, prompting the frogs to either burrow into the mud on the bottom of the pond to minimize dehydration of their bodies and enter a state of estivation, or to attempt to move to a new pond. Despite these behaviors, *X. laevis* can lose as much as 30% of its total body water until rainfall recurs to allow rehydration (Balinsky et al. 1967; Hillman 1978a). This remarkable ability extends the applicability to biological research of an already versatile model organism. Dehydration tolerance by *X. laevis* is facilitated by multiple behavioral, physiological, biochemical, and gene regulatory factors, including shunting nitrogenous waste to urea accumulation (Balinsky et al. 1961, 1967), alteration of enzymatic regulation (Katzenback et al. 2014; Childers and Storey 2016; Dawson et al. 2018), signaling cascades (Malik and Storey 2009a), chaperone proteins (Luu et al. 2018), antioxidants (Malik and Storey 2009b, 2011), and microRNA (Wu et al. 2013; Luu and Storey 2015).

The use of reverse transcription-quantitative polymerase chain reaction (RT-qPCR) is a standard molecular biology technique for measuring gene expression. It is widely used for its sensitivity, high-dynamic range, relative ease-of-use, and cost. Accurate quantification by RT-qPCR requires target values to be normalized to “reference” values, typically one or more genes stably expressed across the experimental conditions of interest. Historically, housekeeping genes such as *GAPDH* and *α-tubulin* were used for normalization as they were assumed to be constantly expressed under most conditions. However, it has become clear that housekeeping genes can often be susceptible to differential expression when authors had assumed otherwise (Selvey et al. 2001; Barber et al. 2005; Greer et al. 2010). Determination of reference genes must therefore be done on a case-by-case basis by identifying genes with stable expression in the specific experimental condition being investigated. There exists a fundamental problem of verifying expression stability: the stability of a reference gene candidate cannot be validated without another reference gene previously determined to be appropriate. This issue has inspired multiple solutions that include popular statistical pipelines incorporated into easy-to-use software packages.

Here we present the validation of reference genes for studying dehydration tolerance in *X. laevis* in eight tissues (brain, dorsal skin, heart, kidney, liver, lung, skeletal muscle, and ventral skin). We use three popular software packages, geNorm (Vandesompele et al. 2002), Normfinder (Andersen et al. 2004), and BestKeeper (Pfaffl et al. 2004) to determine a composite ranking and optimal number of eight genes (*actb.L*, *atp5b1.L*, *canx.L*, *gapdh.S*, *rpl27.L*, *sdha.L*, *tbp.L*, *tbp.S*) to be used for normalization in RT-qPCR experiments.

2.3 Materials and Methods

2.3.1 Animal care and treatment

Animals were treated as previously described (Luu and Storey 2015). Briefly, adult male *X. laevis* from the University of Alberta Science Animal Support Services were acclimated for one week in aerated tanks of dechloraminated water at 22 °C. Animals were fed three times per week, and the water was changed the day after each feeding. Animals were then split into three treatment groups; control, medium dehydration (~15% body water loss), and high dehydration (~30% body water loss) and were not fed again. Animals in the control group were maintained in the above conditions, and medium and high dehydration animals were weighed, and then placed in dry containers at 22 °C to dehydrate. These animals were weighed periodically over several days until the desired percentage water loss was reached. Body water loss was calculated with the following equation:

$$\% \text{ Body water loss} = \frac{m_i - m_d}{m_i \times BWC_i} \times 100\%$$

where m_i is initial mass, m_d is dehydrated mass, and BWC_i is initial body water content (0.74 ± 0.002 g H₂O per g body mass (Malik and Storey 2009a)). Actual body water loss percentages were $16.44 \pm 0.99\%$ and $31.18 \pm 3.20\%$ for medium and high dehydration groups respectively. Animals were then euthanized by pithing and tissues were quickly dissected, flash frozen in liquid nitrogen, and stored at -80 °C until use. All protocols were approved by the Carleton University Animal Care Committee (protocol #106936) and conformed within the guidelines of the Canadian Council on Animal Care.

2.3.2 RNA extraction

RNA was extracted from kidney, lung, dorsal skin, and ventral skin samples from control and high dehydration conditions ($n=6$ per tissue per condition) whereas brain, heart, liver, and skeletal muscle extractions also included the medium dehydration condition ($n=4$ per tissue per condition). For each sample, 50 mg of tissue was homogenized in 1 mL of TRI Reagent (Molecular Research Center) with a Fisher Scientific PowerGen 125 homogenizer. Then, 200 μL of chloroform was added, and the samples were shaken and then centrifuged at $10,000 \times g$ for 15 min at 4 °C. The upper RNA containing aqueous phase was transferred to a new microtube and the above TRI Reagent-chloroform extraction was repeated. A 500 μL aliquot of 2-propanol was then added to each sample to precipitate the RNA at room temperature for 15 min. The samples were centrifuged at $10,000 \times g$ for 15 min at 4 °C, and then the supernatant was discarded. RNA pellets were washed with 70% ethanol, and then air dried and resuspended in 30-50 μL of RNase-free water. RNA concentration and purity were determined with a BioTek Take3 microspot plate and PowerWave HT microplate spectrophotometer. All samples had OD 260/280 ratios of ~ 2.0 and concentrations were normalized to 1 $\mu\text{g}/\mu\text{L}$ with RNase-free water. RNA integrity was then assessed using a 1% agarose gel.

2.3.3 cDNA synthesis

Reverse transcription was performed with 2 μg of RNA from each sample diluted in autoclaved ddH₂O to a volume of 10 μL . A 1 μL aliquot of 200 $\text{ng}\cdot\mu\text{L}^{-1}$ oligo(dT) (Sigma-Aldrich) was added to each tube and tubes were incubated in an Eppendorf Mastercycler thermocycler at 65 °C for 5 min. Samples were cooled on ice and cDNA

synthesis was then performed using the Invitrogen M-MLV Reverse Transcriptase Kit (Cat.# 28025013) with 1 μ L of 10 mM dNTPs (BioShop, Burlington, Canada). Samples were then stored at -20 °C until use.

2.3.4 Candidate reference gene selection and primer design

Candidate reference genes were selected using one of three criteria: i) traditional housekeeping genes (*actb.L*, *gapdh.S*), ii) reference genes previously reported as suitable in similar systems (*atp5bfl.L*, *tbp.L*, *tbp.S*, *canx.L* (Benn et al. 2008; Turabelidze et al. 2010; Lanzafame et al. 2015)), or iii) genes showing high expression levels in multiple adult *X. laevis* tissues as shown in previously published transcriptomes (*rpl27.L*, *sdha.L*; NCBI GEO: GSE73419). Primers were designed using Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) with PCR product size set to 80-120 bp (unless otherwise specified), primer pair specificity checking for *Xenopus laevis* (taxid: 8355), and the remaining options left as default. Primers were synthesized by Integrated DNA Technologies and sequences are shown in **Table 2.1**.

2.3.5 RT-qPCR

RT-qPCR was performed using a Bio-Rad CFX Connect as previously described (Pellissier et al. 2006). Reactions contained 2 μ L cDNA, 2 μ L qPCR buffer (100 mM Tris-HCl [pH 8.5], 500 mM KCl, 1.5% v:v Triton X-100, 20 mM MgCl₂, and 1 mM fluorescein), 0.16 μ L of 25 mM dNTPs, 4 μ L of 1 M trehalose, 0.5 μ L of formamide, 0.1 μ L of 100X SYBR Green in DMSO (Invitrogen; #S7585), 0.5 μ L of 0.3 nmol· μ L⁻¹ of each primer, 0.125 μ L of 5U· μ L⁻¹ wild-type Taq (BioShop; #TAQ001.1) and autoclaved ddH₂O for a final volume of 20 μ L. A 2-fold dilution series standard curve of pooled samples was run with each quantification run.

Thermocycling conditions consisted of an initial denaturing step at 95 °C for 3 min, then 50 cycles of: denaturing at 95 °C for 10 sec, primer annealing at 57 °C for 20 sec, and extension at 72 °C for 20 sec. Post-run melt analyses from 55-95 °C were performed and serial 2-fold dilution standard curves with a minimum of five points for each target were run on quantification plates to determine reaction efficiencies. All primer-pairs yielded a single peak on melt-curve derivative plots, a product at the expected size when run on a 3% agarose gel (**Supplementary Figure 2.1**) and had efficiencies between 90-110%. The dilution factor of the cDNA for each tissue was optimized based on the standard curves. For results comparing tissues directly, C_q values obtained from CFX Manager (BioRad) were adjusted to account for varying dilution factors between tissues (**Figure 2.1, Figure 2.2**).

2.3.6 Reference gene stability analysis and validation

To analyze the stability of the candidate reference genes, we employed three software packages: geNorm (Vandesompele et al. 2002), Normfinder (Andersen et al. 2004), and BestKeeper (Pfaffl et al. 2004). For geNorm and Normfinder, C_q values were transformed to relative quantities and used as input using the Pfaffl method (Pfaffl 2001):

$$Q = (E\% * 0.01 + 1)^{C_{qmin} - C_{qsample}}$$

where Q is the relative quantity, $E\%$ is the efficiency percentage of the PCR reaction, C_{qmin} is the lowest C_q of that gene, and $C_{qsample}$ is the C_q of the sample. For BestKeeper, C_q values are required as input, thus, to account for true PCR reaction efficiencies, C_q values obtained from CFX Manager were converted first to relative quantities as described above, then to corrected C_q values using the following equation:

$$C_{qcorr} = C_{qmin} - \log_2 Q$$

where C_{qcorr} is the corrected C_q , C_{qmin} is the lowest C_q of that gene, and Q is the relative quantity. Stability values for all eight genes from each tissue were obtained from all three software packages and were assigned a rank from most stable (rank 1) to least stable (rank 8). A composite rank for each gene in each tissue was determined by averaging the ranks from each software package.

2.3.7 Determination of optimal number of reference genes

To determine the optimal number of reference genes to use for normalization for each tissue, the geNorm pairwise variation (Vandesompele et al. 2002) method was adapted to use the composite rankings instead of the rankings solely produced from the geNorm software. A normalization factor ($NF_{n,i}$) for the top n genes of sample i is calculated using the geometric mean of relative quantities (Q) for each gene for that sample:

$$NF_{n,i} = \sqrt[n]{Q_{1,i} * Q_{2,i} * \dots * Q_{n,i}}$$

An array ($A_{n/n+1}$) of the \log_2 transformed ratios of $NF_{n,i}$ and $NF_{n+1,i}$ was calculated:

$$A_{n/n+1} = \left\{ \log_2 \frac{NF_{n,1}}{NF_{n+1,1}}, \log_2 \frac{NF_{n,2}}{NF_{n+1,2}}, \dots, \log_2 \frac{NF_{n,i}}{NF_{n+1,i}} \right\}$$

The pairwise variation of the top n and $n+1$ genes ($V_{n/n+1}$) was then determined by taking the SD of $A_{n/n+1}$. The pairwise variation is a measure of the effect of adding the next most stable gene to the calculation of the normalization factor. The recommended cut-off of 0.15 as previously established (Vandesompele et al. 2002), was used to determine the optimal number of reference genes. A $V_{n/n+1}$ under this value indicates that adding the $n+1^{\text{th}}$ gene has little effect on the normalization factors. For example, if $V_{2/3} < 0.15$, two reference genes would be determined to be optimal, whereas if $V_{2/3} > 0.15$ and $V_{3/4} < 0.15$, then three reference genes would be optimal.

2.4 Results

2.4.1 Expression levels of candidate reference genes

All primer pairs amplified a single product at the expected size (**Supplementary Figure 2.1, Table 2.1**). All genes in all samples had quantification cycle (C_q) values between 20 and 38 (**Supplementary Table 2.1**). The C_q values of each gene varied widely across all samples, suggesting that no single gene would be a suitable reference gene for every tissue (**Figure 2.1**). The gene *actb.L* had the highest expression across all tissues (median $C_q = 26.32$), followed by *rpl27.L* (27.25), *gapdh.S* (28.08), *sdha.L* (30.48), *canx.L* (31.76), *tbp.S* (33.03), *tbp.L* (33.29), and *atp5f1b.L* (34.08). When C_q values were separated by tissue, experimental condition, and gene (**Figure 2.2**), potentially stably expressed genes became apparent but varied by tissue. For example, *canx.L* had constant expression between experimental conditions in brain, dorsal skin, skeletal muscle, and ventral skin, but varied by experimental condition in heart, kidney, liver, and lung.

2.4.2 GeNorm Analysis

Tissues were analyzed separately with geNorm, and stability (M) values are in **Supplementary Table 2.2**. Smaller M values indicate greater stability, and genes were ranked from most (rank 1) to least (rank 8) stable by tissue type. The geNorm algorithm does not distinguish between the two most stable genes, thus, the two most stable genes for each tissue have the same rank and M value. Top ranked genes varied by tissue; *atp5f1b.L* and *tbp.L* (M -value = 0.201) in brain, *gapdh.S* and *sdha.L* (0.186) in dorsal skin, *gapdh.S* and *rpl27.L* (0.424) in heart, *actb.L* and *gapdh.S* (0.564) in kidney, *canx.L* and *sdha.L* (0.314) in liver, *actb.L* and *gapdh.S* (0.189) in lung, *gapdh.S* and *rpl27.L*

(0.265) in muscle, and *rpl27.L* and *tbp.L* (0.277) in ventral skin. In every tissue assessed, all genes were below the recommended 1.5 *M* threshold value (Vandesompele et al. 2002).

2.4.3 Normfinder Analysis

Tissues were analyzed separately with Normfinder as well, and stability values are in **Supplementary Table 2.3**. The Normfinder algorithm differs from geNorm as it incorporates information about groups (experimental conditions), and determines stabilities based on intergroup and intragroup differences. Each gene was again ranked from most to least stable, and interestingly, the rankings determined by Normfinder differed from the geNorm rankings. By tissues, the most stable genes were *canx.L* (Stability value = 0.06) in brain, *rpl27.L* (0.11) in dorsal skin, *actb.L* (0.21) in heart, *actb.L* (0.20) in kidney, *atp5fb1.L* (0.18) in liver, *canx.L* (0.3) in lung, *rpl27.L* (0.19) in muscle, and *canx.L* (0.14) in ventral skin.

2.4.4 BestKeeper Analysis

We used the BestKeeper algorithm to add a third measure of stability of the candidate reference genes in each tissue, and the rankings are in **Supplementary Table 2.4**. The web application that hosts BestKeeper (<http://leonxie.esy.es/RefFinder/>) uses the standard deviation (SD) of C_q values for each gene as a measure of stability. The gene with the lowest SD is thus ranked highest (rank 1) while the gene with the highest SD is ranked lowest (rank 8). The rankings produced from BestKeeper also differed from the previous two algorithms, where *canx.L* (SD=0.15) ranked highest in brain, *rpl27.L* (0.17) in dorsal skin, *actb.L* (0.34) in heart, *rpl27.L* (0.48) in kidney, *rpl27.L* (0.15) in liver, *rpl27.L* (0.09) in lung, *gapdh.S* (0.13) in muscle, and *rpl27.L* (0.18) in ventral skin.

2.4.5 Candidate reference gene composite ranking

We performed a composite ranking for each tissue by averaging the rank of each gene from the three algorithms (**Figure 2.3**). No single gene had the highest composite rank in all tissues, however, some genes tended to rank higher on average (*rpl27.L*), while others tended to rank lower (*tbp.S*). Interestingly, the ranking of the three algorithms were highly similar in tissues such as dorsal and ventral skin, heart, and muscle, whereas the rankings in brain and liver varied more between the algorithms. The highest composite ranked genes were *canx.L* (composite rank = 2.33) in brain, *rpl27.L* (1.67) in dorsal skin, *actb.L* (1.67) in heart, *actb.L* (1.67) in kidney, *sdha.L* (2.67) in liver, *actb.L* (3) in lung, *gapdh.S* (1.33) in muscle, and *rpl27.L* (1.33) in ventral skin.

2.4.6 Optimal number of reference genes for each tissue

To determine how many reference genes should be used we adapted the pairwise-variation method from the geNorm algorithm to use the composite rankings produced from all three algorithms (**Figure 2.4**). This method starts by calculating the normalization factors for the top two ranked genes in each tissue, then determines how much the normalization factors change when the next highest ranked gene is added. This is repeated until all eight genes are included. We found that the variation produced when adding a third reference gene was not greater than the threshold cut-off value (0.15) in all tissues except kidney. Thus, the top two ranked reference genes in each tissue was deemed sufficient, with the exception of kidney, where the use of more reference genes may be beneficial. In kidney, pairwise variations were above the threshold cut-off value until the seventh gene was added, suggesting that the top six genes be used for reference

gene normalization. **Table 2.2** incorporates both the composite rankings and optimal reference gene number for each tissue where the optimal genes are bold.

2.5 Discussion

RT-qPCR is a powerful and sensitive technique that relies on stringent validation of an internal control to account for variable efficiencies in RNA extraction (El-Khoury et al. 2016), cDNA conversion (Bustin and Nolan 2004), and experimental error included from technical and human sources during sample preparation and treatments to ensure reproducibility (Bustin et al. 2009). The most common internal control strategy is to use a stably expressed gene; one that does not show regulation between experimental conditions since this gene and the target genes will undergo all sample processing steps together. In fact, this is the only method for normalization recommended in the MIQE guidelines (Bustin et al. 2009). The determination of a stably expressed gene, otherwise known as a reference gene, presents a circular challenge as this gene also needs to be standardized to an internal control. As a result multiple algorithmic strategies have been used to validate and rank candidate reference genes; the three most common being geNorm (Vandesompele et al. 2002), Normfinder (Andersen et al. 2004), and BestKeeper (Pfaffl et al. 2004).

Validation of reference genes has been done in many other systems (Turabelidze et al. 2010; Wang et al. 2012, 2015b, a; Imai et al. 2014; Lanzafame et al. 2015; Machado et al. 2015; Shivhare and Lata 2016) including developing *X. laevis* (Mughal et al. 2018), however a comprehensive assessment of reference genes in adult *X. laevis* with respect to dehydration tolerance has yet to be performed. Here, we present a survey of eight

candidate reference genes (**Table 2.1**) in eight tissues of *X. laevis* subjected to ~15% or ~30% dehydration.

2.5.1 Candidate reference gene expression varies by tissue

Determining suitable reference genes in this system is particularly difficult due to the varied physiological and molecular responses that each tissue elicits in response to dehydration. For example, dorsal and ventral skin synthesize and secrete a mucous coating to minimize evaporative water loss (Hillman 1978b). Water is preferentially lost from plasma, thus, increasing blood viscosity (Hillman 1978b) and prompting an increased cardiac output to maintain oxygen delivery (Hillman 1978a). The liver synthesizes and secretes urea, both as a reservoir for nitrogenous waste when ammonia can no longer be secreted into the aquatic environment and potentially to increase the osmotic pressure of body fluids to decrease water loss (Malik and Storey 2009a). Blood circulation is reduced to the skeletal muscles, whereas it is elevated in the brain (Hillman and Sommerfeldt 1981), and the kidneys must contend with a many-fold increase in various plasma solutes (Hillman 1978b, a). The molecular responses to whole-body dehydration implicates many metabolic and structural processes, and as a result, we believed that further steps were required to confirm the suitability of the reference genes for each tissue that was investigated. This is demonstrated in **Figure 2.1** where there is a wide range of C_q values across all samples for each gene. When the candidate reference genes are separated by tissue and experimental condition (**Figure 2.2**), the amount of variation exhibited by these genes also varied between tissues. For example, the candidate reference genes in brain appear more constant in general than in liver or lung. The response to dehydration of individual genes also appears to vary by tissue. In brain,

kidney, and ventral skin, *atp5f1b.L* was constantly expressed between conditions, but showed considerable regulation in dorsal skin, heart, liver, lung, and skeletal muscle.

2.5.2 Composite ranking show tissue-specific order

The three algorithms used employ different strategies for ranking the stability of reference genes. The geNorm algorithm's measure of stability (*M*-value) for a gene is the arithmetic mean of all pairwise variations of normalization factors with the other genes being analyzed. Normfinder incorporates experimental group membership in a model that estimates a stability value using inter- and intra-group variation. Thirdly, for the purpose of ranking reference genes, BestKeeper simply calculates gene stability based on SD. Unsurprisingly, these methods produced ranking results that varied to different degrees by tissue (**Figure 2.3**). The algorithms converged on rankings in dorsal and ventral skin, muscle, and heart, whereas there was much less convergence in brain, kidney, liver, and lungs. As discussed above, the physiological and molecular responses to dehydration employed by these tissues suggest that no single gene would be an optimal internal control for all tissues, and this was reflected in our results. To produce a single rank for each gene, we created composite rankings by averaging the ranks from the three algorithms. While no single gene was ranked highest for all tissues using this method, some genes tended to rank higher across tissues than others. For example, *rpl27.L* was ranked in the top two in all tissues except kidney and liver, whereas *atp5f1b.L* was ranked in the top two only in liver. Interestingly, although the algorithms produced congruent rankings in dorsal and ventral skin, the rankings they produced differed between these highly similar tissues, suggesting differing responses to dehydration.

2.5.3 Two reference genes are sufficient in most tissues

The use of more than one reference gene for normalization is a common strategy since it can be difficult to find a single suitable gene, whereas the use of multiple genes may provide more reliable results. We employed a pairwise variation analysis strategy (Vandesompele et al. 2002) to determine the optimal number of reference genes to use (Figure 2.4). The top two compositely ranked genes in all tissues was deemed optimal except for kidney which required the top six. These values are combined with the gene rankings in Table 2.2 for our determination of tissue specific reference gene sets (bold) for use in *X. laevis* dehydration tolerance studies.

2.5.4 Conclusions

We have identified tissue specific reference gene sets for normalization in studies on dehydration tolerance in *X. laevis*. A common set of eight candidate reference genes (*actb.L*, *atp5bf1.L*, *canx.L*, *gapdh.S*, *rpl27.L*, *sdha.L*, *tbp.L*, *tbp.S*) was screened in eight tissues (brain, dorsal skin, heart, kidney, liver, lung, skeletal muscle, and ventral skin) using three popular algorithms (geNorm, Normfinder, and BestKeeper). Reference gene stability rankings from each algorithm were combined to form a composite rank in each tissue, and the optimal number of reference genes was evaluated using a pairwise variation method. The following gene sets were determined for each tissue: *canx.L* and *rpl27.L* in brain, *rpl27.L* and *sdha.L* in dorsal skin, *actb.L* and *rpl27.L* in heart, *actb.L*, *gapdh.S*, *tbp.L*, *atp5f1b.L* *tbp.S* and *rpl27.L* in kidney, *sdha.L* and *atp5f1b.L* in liver, *actb.L* and *rpl27.L* in lung, *gapdh.S* and *rpl27.L* in skeletal muscle, and *rpl27.L* and *canx.L* in ventral skin. The results of the present study recommend that these reference gene sets be used for normalization of gene quantification in future studies of stress

responses to dehydration in *X. laevis*, and potentially to analyze gene responses to other environmental stresses. While it is unlikely a single gene is stably expressed in all tissues, future studies could incorporate next-generation sequencing technologies to screen a high number of genes and determine whether a common set of genes are stable in all tissues.

2.6 References

- Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–50 . doi: 10.1158/0008-5472.CAN-04-0496
- Balinsky JB, Choritz EL, Coe CG, van der Schans GS (1967) Amino acid metabolism and urea synthesis in naturally aestivating *Xenopus laevis*. *Comp Biochem Physiol* 22:59–68 . doi: 10.1016/0010-406x(67)90166-1
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Barber RD, Harmer DW, Coleman RA, Clark BJ (2005) GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics* 21:389–95 . doi: 10.1152/physiolgenomics.00025.2005
- Benn CL, Fox H, Bates GP (2008) Optimisation of region-specific reference gene selection and relative gene expression analysis methods for pre-clinical trials of Huntington's disease. *Mol Neurodegener* 3:17 . doi: 10.1186/1750-1326-3-17
- Bustin SA, Benes V, Garson JA, et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–22 . doi: 10.1373/clinchem.2008.112797
- Bustin SA, Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 15:155–66
- Childers CL, Storey KB (2016) Post-translational Regulation of Hexokinase Function and Protein Stability in the Aestivating Frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- El-Khoury V, Pierson S, Kaoma T, et al (2016) Assessing cellular and circulating miRNA recovery: the impact of the RNA isolation method and the quantity of input material. *Sci Rep* 6:19529 . doi: 10.1038/srep19529
- Greer S, Honeywell R, Geletu M, et al (2010) Housekeeping genes; expression levels may change with density of cultured cells. *J Immunol Methods* 355:76–9 . doi: 10.1016/j.jim.2010.02.006
- Gurdon JB (1962) Adult frogs derived from the nuclei of single somatic cells. *Dev Biol* 4:256–73 . doi: 10.1016/0012-1606(62)90043-X
- Hillman SS (1978a) The roles of oxygen delivery and electrolyte levels in the dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Hillman SS (1978b) Some effects of dehydration on internal distributions of water and solutes in *Xenopus laevis*. *Comp Biochem Physiol Part A Physiol* 61:303–307

- Hillman SS, Sommerfeldt RW (1981) Microsphere studies of amphibian systemic blood flow redistribution during dehydration, hypovolemia, and salt load. *J Exp Zool* 218:305–308
- Imai T, Ubi BE, Saito T, Moriguchi T (2014) Evaluation of reference genes for accurate normalization of gene expression for real time-quantitative PCR in *Pyrus pyrifolia* using different tissue samples and seasonal conditions. *PLoS One* 9:e86492 . doi: 10.1371/journal.pone.0086492
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Lanzafame M, Botta E, Teson M, et al (2015) Reference genes for gene expression analysis in proliferating and differentiating human keratinocytes. *Exp Dermatol* 24:314–6 . doi: 10.1111/exd.12657
- Luu BE, Storey KB (2015) Dehydration triggers differential microRNA expression in *Xenopus laevis* brain. *Gene* 573:64–9
- Luu BE, Wijenayake S, Malik AI, Storey KB (2018) The regulation of heat shock proteins in response to dehydration in *Xenopus laevis*. *Cell Stress Chaperones* 23:45–53 . doi: 10.1007/s12192-017-0822-9
- Machado RD, Christoff AP, Loss-Morais G, et al (2015) Comprehensive selection of reference genes for quantitative gene expression analysis during seed development in *Brassica napus*. *Plant Cell Rep* 34:1139–49 . doi: 10.1007/s00299-015-1773-1
- Malik AI, Storey KB (2009a) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Malik AI, Storey KB (2011) Transcriptional regulation of antioxidant enzymes by FoxO1 under dehydration stress. *Gene* 485:114–9 . doi: 10.1016/j.gene.2011.06.014
- Malik AI, Storey KB (2009b) Activation of antioxidant defense during dehydration stress in the African clawed frog. *Gene* 442:99–107 . doi: 10.1016/j.gene.2009.04.007
- Mughal BB, Leemans M, Spirhanzlova P, et al (2018) Reference gene identification and validation for quantitative real-time PCR studies in developing *Xenopus laevis*. *Sci Rep* 8:496 . doi: 10.1038/s41598-017-18684-1
- Pellissier F, Glogowski CM, Heinemann SF, et al (2006) Lab assembly of a low-cost, robust SYBR green buffer system for quantitative real-time polymerase chain reaction. *Anal Biochem* 350:310–2 . doi: 10.1016/j.ab.2005.12.002
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45 . doi: 10.1093/nar/29.9.e45
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26:509–15 . doi: 10.1023/B:BILE.0000019559.84305.47
- Selvey S, Thompson EW, Matthaei K, et al (2001) Beta-actin--an unsuitable internal

- control for RT-PCR. *Mol Cell Probes* 15:307–11 . doi: 10.1006/mcpr.2001.0376
- Shivhare R, Lata C (2016) Selection of suitable reference genes for assessing gene expression in pearl millet under different abiotic stresses and their combinations. *Sci Rep* 6:23036 . doi: 10.1038/srep23036
- Turabelidze A, Guo S, DiPietro LA (2010) Importance of housekeeping gene selection for accurate reverse transcription-quantitative polymerase chain reaction in a wound healing model. *Wound Repair Regen* 18:460–6 . doi: 10.1111/j.1524-475X.2010.00611.x
- Vandesompele J, De Preter K, Pattyn F, et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034
- Wang E, Wang K, Chen D, et al (2015a) Evaluation and Selection of Appropriate Reference Genes for Real-Time Quantitative PCR Analysis of Gene Expression in Nile Tilapia (*Oreochromis niloticus*) during Vaccination and Infection. *Int J Mol Sci* 16:9998–10015 . doi: 10.3390/ijms16059998
- Wang H, Chen S, Jiang J, et al (2015b) Reference gene selection for cross-species and cross-ploidy level comparisons in *Chrysanthemum spp.* *Sci Rep* 5:8094 . doi: 10.1038/srep08094
- Wang Q, Ishikawa T, Michiue T, et al (2012) Stability of endogenous reference genes in postmortem human brains for normalization of quantitative real-time PCR data: comprehensive evaluation using geNorm, NormFinder, and BestKeeper. *Int J Legal Med* 126:943–52 . doi: 10.1007/s00414-012-0774-7
- Wu CW, Biggar KK, Storey KB (2013) Dehydration mediated microRNA response in the African clawed frog *Xenopus laevis*. *Gene* 529:269–75

Table 2.1 Candidate reference gene primers used for qPCR analysis in *X. laevis*. Each primer pair was used for all eight tissues of interest. NCBI RefSeq accession numbers are provided when available. The Xenbase accession number is provided for *atp5f1b.L*

Gene symbol	Full name	Accession	Forward Primer	Reverse Primer	Product Size
<i>actb.L</i>	Beta-actin	NM_001088953	GACAGTCTGTGTGCGTCCAA	TGGGCGACCCACAATAGATG	147
<i>atp5f1b.L</i>	ATP synthase F1 subunit beta	17340769	TGCCCGTTTGTCTTGGGTAA	AACAGCCACACAAGTTGGTC	83
<i>canx.L</i>	Calnexin	NM_001086857	TACTCCCTGTTCCACACAAGT	AGTGCTGGCAGCTTAACAGT	112
<i>gapdh.S</i>	Glyceraldehyde-3-phosphate dehydrogenase	NM_001087098	GCCGCCATTAAGACTGCATCA	TGGAGACAACCTGGTCTTGTG	73
<i>rpl27.L</i>	Ribosomal protein L27	NM_001086954	GAGTTGGAGTGAAAATGGGCA	TCACAATGACAGCTTTGCGG	93
<i>sdha.L</i>	Succinate dehydrogenase complex subunit A	NM_001090004	TCACGGCTTCTTAGCAGAGC	TGCTGAAGTGTCTTCCGTC	120
<i>tbp.L</i>	TATA-box binding protein	NM_001090900	GTGGGGCTTTGTTTCCTGATG	AGTTCCTGTTGACCAGCTTTC	114
<i>tbp.S</i>	TATA-box binding protein	NM_001095569	ACAGTGGCAGAAGACCCCTA	GTCCTCCATGAAACAAAGCCC	81

Table 2.2 Composite rankings for reference gene candidates in eight tissue of *X. laevis*. Candidate reference genes are first ranked by stability using geNorm, Normfinder, and BestKeeper software. These ranks are averaged for each candidate gene in each tissue to get a composite rank. Composite ranks are in parenthesis and top *N* genes as determined in **Figure 2.4** are bold

Rank	Brain	Dorsal	Heart	Kidney	Liver	Lung	Muscle	Ventral
1	<i>canx.L (2.33)</i>	<i>rpl27.L (1.67)</i>	<i>actb.L (1.67)</i>	<i>actb.L (1.67)</i>	<i>sdha.L (2.67)</i>	<i>actb.L (3)</i>	<i>gapdh.S (1.33)</i>	<i>rpl27.L (1.33)</i>
2	<i>rpl27.L (2.67)</i>	<i>sdha.L (2.33)</i>	<i>rpl27.L (2)</i>	<i>gapdh.S (1.67)</i>	<i>atp5f1b.L (3)</i>	<i>rpl27.L (3)</i>	<i>rpl27.L (1.67)</i>	<i>canx.L (2.33)</i>
3	<i>atp5f1b.L (4)</i>	<i>canx.L (3.33)</i>	<i>gapdh.S (2.33)</i>	<i>tbp.L (3.33)</i>	<i>canx.L (3)</i>	<i>gapdh.S (3.67)</i>	<i>actb.L (3.67)</i>	<i>tbp.L (2.67)</i>
4	<i>gapdh.S (4.33)</i>	<i>gapdh.S (3.33)</i>	<i>canx.L (4.67)</i>	<i>atp5f1b.L (5.33)</i>	<i>actb.L (3.67)</i>	<i>sdha.L (3.67)</i>	<i>canx.L (3.67)</i>	<i>sdha.L (3.67)</i>
5	<i>actb.L (4.67)</i>	<i>actb.L (4)</i>	<i>tbp.L (5)</i>	<i>tbp.S (5.33)</i>	<i>gapdh.S (4)</i>	<i>canx.L (4)</i>	<i>tbp.S (5)</i>	<i>tbp.S (4.67)</i>
6	<i>tbp.L (4.67)</i>	<i>tbp.S (6.33)</i>	<i>atp5f1b.L (5.67)</i>	<i>rpl27.L (5.67)</i>	<i>rpl27.L (5)</i>	<i>tbp.L (4.33)</i>	<i>tbp.L (5.33)</i>	<i>actb.L (6.67)</i>
7	<i>tbp.S (5)</i>	<i>atp5f1b.L (6.67)</i>	<i>tbp.S (6.67)</i>	<i>sdha.L (5.67)</i>	<i>tbp.L (6.67)</i>	<i>tbp.S (6)</i>	<i>sdha.L (7)</i>	<i>gapdh.S (7)</i>
8	<i>sdha.L (8)</i>	<i>tbp.L (8)</i>	<i>sdha.L (7.67)</i>	<i>canx.L (7)</i>	<i>tbp.S (7.67)</i>	<i>atp5f1b.L (8)</i>	<i>atp5f1b.L (8)</i>	<i>atp5f1b.L (7.33)</i>

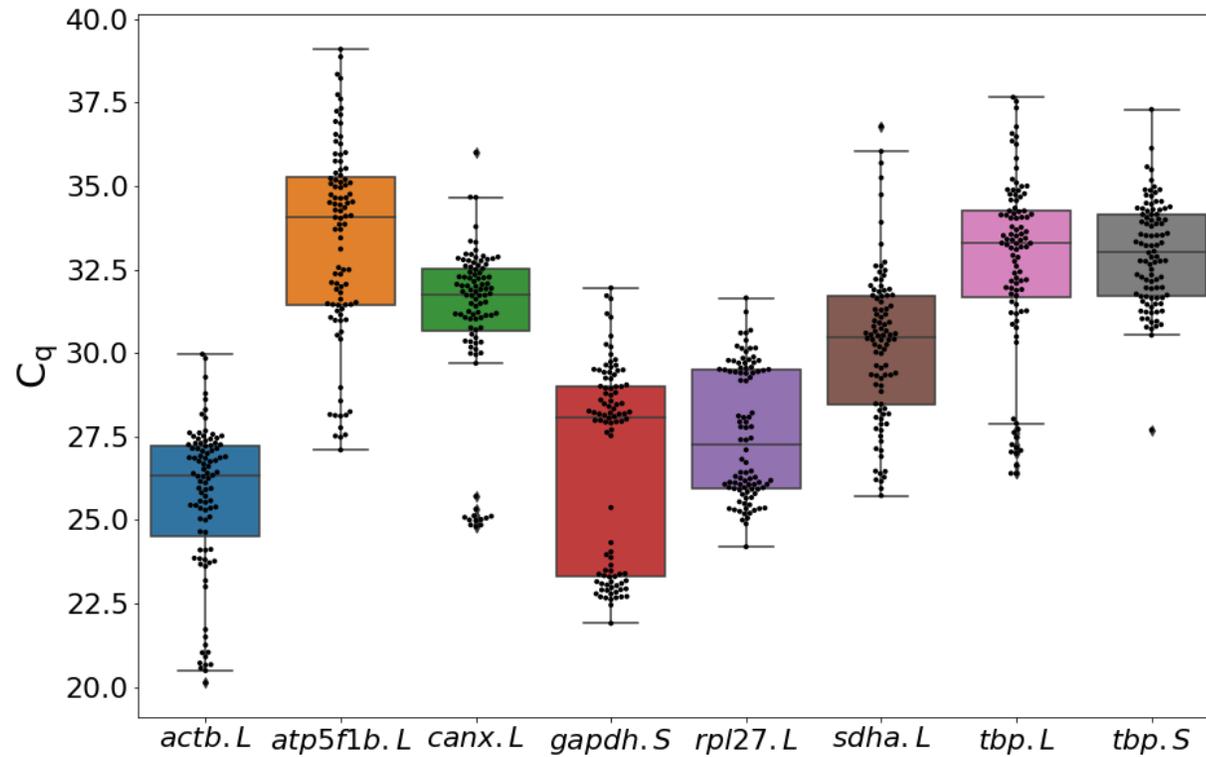


Figure 2.1 Distribution of quantification cycles of all samples for each candidate reference gene. Boxes extend from the first to third quartile where the center horizontal line is the second quartile. Whiskers extend to the 5th and 95th percentiles

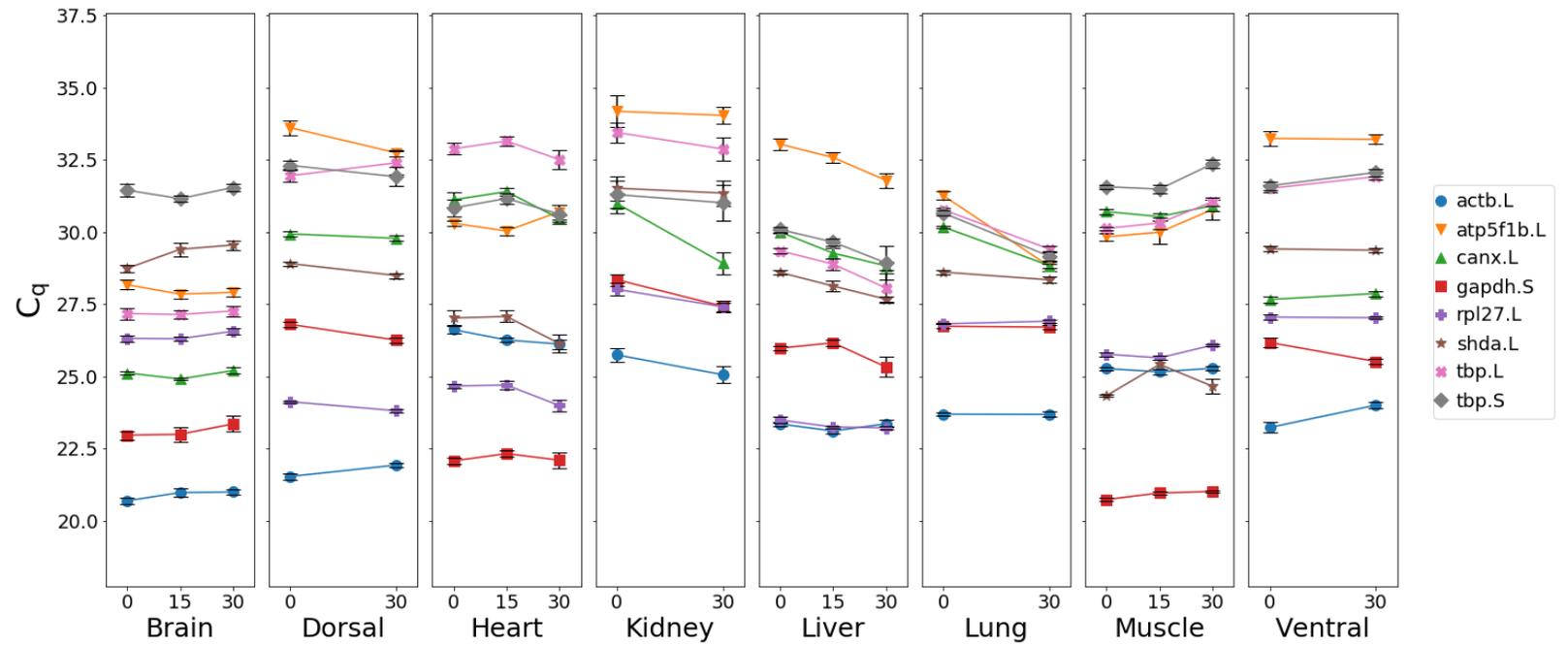


Figure 2.2 Mean quantification cycle for each experimental condition of each tissue. Data are mean \pm SE, $n=4-6$ depending on tissue

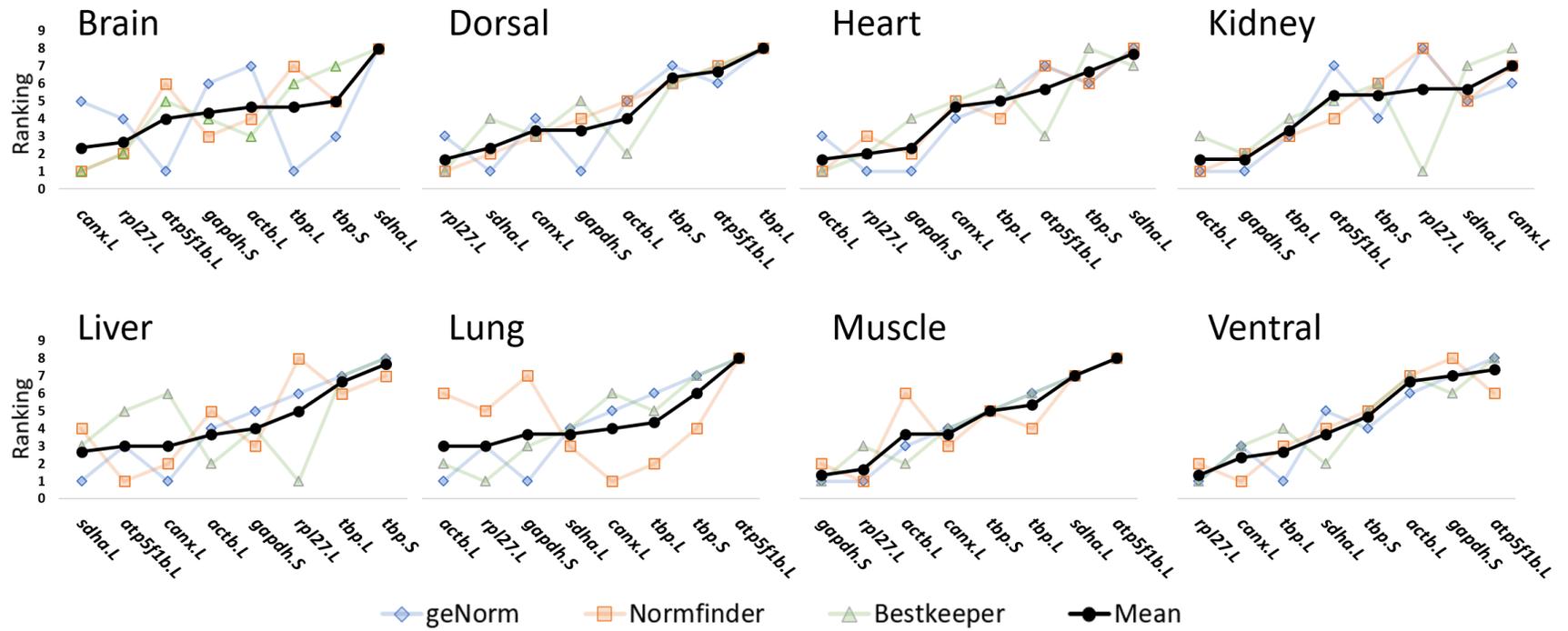


Figure 2.3 Composite and component rankings for reference gene candidates in eight tissue of *X. laevis*. Lower ranks indicate better stability and suitability as a reference gene. Component stabilities are produced by (◆) geNorm, (■) Normfinder, (▲) BestKeeper. The composite ranking (●) is the mean of all component rankings

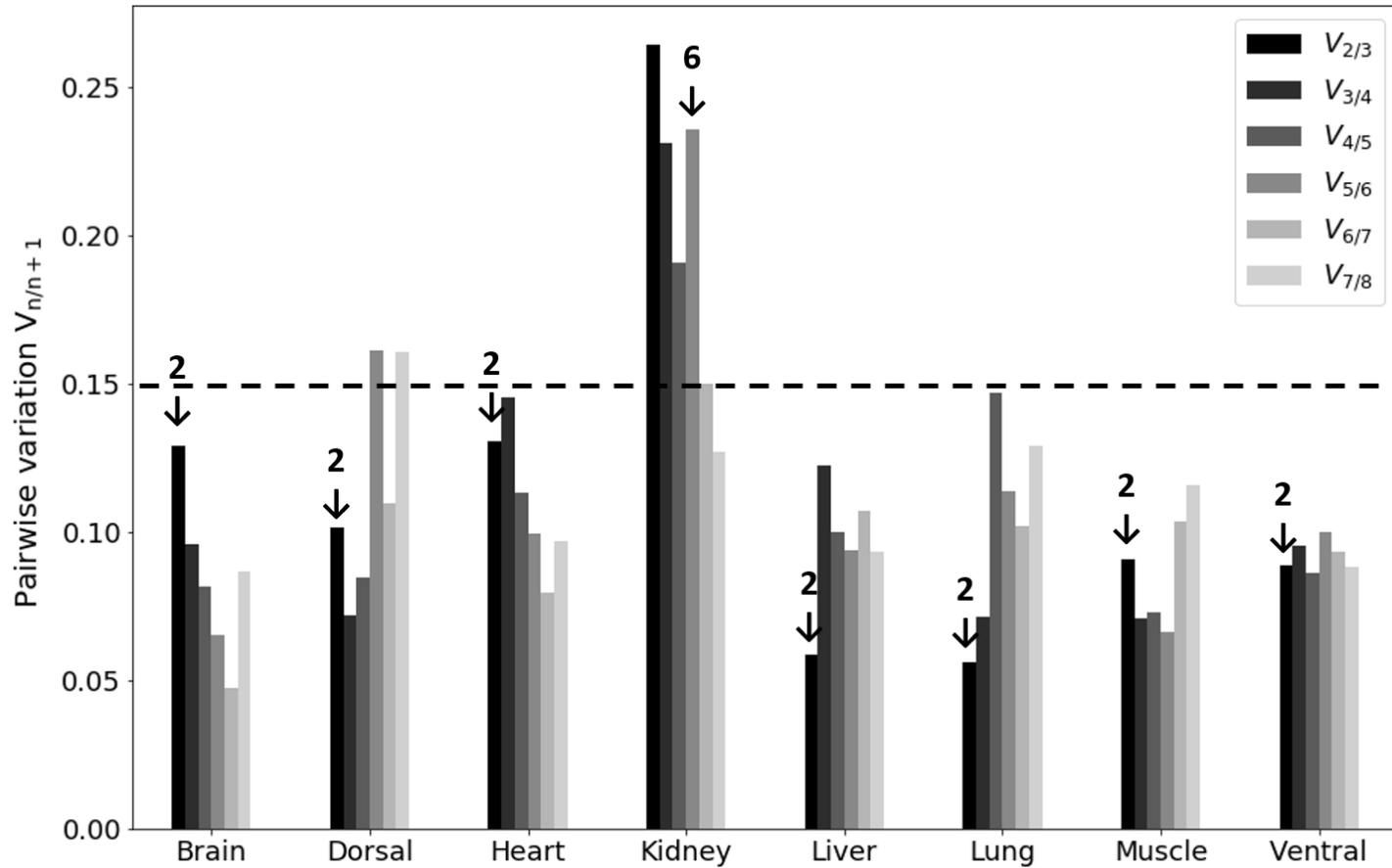


Figure 2.4 Determination of optimal number of reference genes. Pairwise variation (V) of n and $n+1$ gene containing normalization factors (NF_n, NF_{n+1}). Higher V indicate higher effect of adding the next most stable reference gene by composite rank (Table 2.2) for normalization. The recommend cut-off of 0.15 from (Vandesompele et al. 2002) was used to determine when a subsequent reference gene should be included. The optimal number is indicated for each tissue

Supplementary Table 2.1 C_q values for brain, heart, muscle, and liver of 0%, 15%, and 30% dehydrated *Xenopus laevis* (mean ± SE; n=4 per condition), and kidney, dorsal skin, ventral skin, and lung of 0% and 30% dehydrated *X. laevis* (mean ± SE; n=6 per condition)

Tissue	Condition	<i>actb.L</i>	<i>atp5f1b.L</i>	<i>canx.L</i>	<i>gapdh.S</i>	<i>rpl27.L</i>	<i>shda.L</i>	<i>tbp.L</i>	<i>tbp.S</i>
Brain	0%	20.69 ± 0.12	28.17 ± 0.16	25.12 ± 0.05	22.96 ± 0.15	26.32 ± 0.10	28.74 ± 0.14	27.17 ± 0.19	31.45 ± 0.22
Brain	15%	20.98 ± 0.15	27.85 ± 0.14	24.91 ± 0.04	22.99 ± 0.25	26.30 ± 0.08	29.40 ± 0.24	27.14 ± 0.14	31.16 ± 0.12
Brain	30%	21.00 ± 0.10	27.91 ± 0.14	25.20 ± 0.10	23.35 ± 0.28	26.57 ± 0.10	29.55 ± 0.16	27.27 ± 0.18	31.54 ± 0.14
Heart	0%	26.62 ± 0.14	30.30 ± 0.10	31.12 ± 0.26	22.08 ± 0.10	24.67 ± 0.07	27.02 ± 0.26	32.88 ± 0.19	30.84 ± 0.30
Heart	15%	26.26 ± 0.06	30.04 ± 0.15	31.40 ± 0.11	22.32 ± 0.10	24.70 ± 0.13	27.07 ± 0.20	33.15 ± 0.16	31.16 ± 0.17
Heart	30%	26.12 ± 0.16	30.70 ± 0.26	30.46 ± 0.10	22.10 ± 0.28	23.99 ± 0.19	26.13 ± 0.31	32.50 ± 0.32	30.61 ± 0.31
Muscle	0%	25.27 ± 0.06	29.83 ± 0.12	30.70 ± 0.08	20.73 ± 0.06	25.77 ± 0.08	24.33 ± 0.04	30.13 ± 0.06	31.57 ± 0.08
Muscle	15%	25.16 ± 0.09	29.99 ± 0.39	30.53 ± 0.12	20.96 ± 0.04	25.64 ± 0.07	25.41 ± 0.15	30.31 ± 0.19	31.49 ± 0.13
Muscle	30%	25.28 ± 0.07	30.78 ± 0.36	30.92 ± 0.19	21.01 ± 0.04	26.08 ± 0.04	24.66 ± 0.26	31.05 ± 0.15	32.36 ± 0.13
Liver	0%	23.35 ± 0.08	33.04 ± 0.19	30.00 ± 0.06	25.98 ± 0.06	23.49 ± 0.10	28.60 ± 0.07	29.35 ± 0.09	30.08 ± 0.09
Liver	15%	23.11 ± 0.10	32.58 ± 0.18	29.27 ± 0.19	26.16 ± 0.12	23.25 ± 0.05	28.13 ± 0.18	28.89 ± 0.21	29.65 ± 0.13
Liver	30%	23.36 ± 0.14	31.79 ± 0.26	28.83 ± 0.15	25.32 ± 0.34	23.22 ± 0.07	27.67 ± 0.09	28.06 ± 0.50	28.93 ± 0.57
Kidney	0%	25.74 ± 0.24	34.18 ± 0.55	30.99 ± 0.34	28.35 ± 0.20	28.02 ± 0.22	31.52 ± 0.43	33.45 ± 0.35	31.30 ± 0.48
Kidney	30%	25.06 ± 0.29	34.04 ± 0.31	28.92 ± 0.39	27.43 ± 0.21	27.41 ± 0.14	31.35 ± 0.44	32.88 ± 0.40	31.01 ± 0.63
Dorsal	0%	21.53 ± 0.11	33.61 ± 0.26	29.93 ± 0.08	26.80 ± 0.08	24.12 ± 0.04	28.91 ± 0.07	31.95 ± 0.22	32.31 ± 0.17
Dorsal	30%	21.93 ± 0.08	32.74 ± 0.11	29.78 ± 0.09	26.25 ± 0.08	23.81 ± 0.05	28.49 ± 0.09	32.40 ± 0.42	31.91 ± 0.32
Ventral	0%	23.23 ± 0.17	33.24 ± 0.25	27.66 ± 0.11	26.17 ± 0.15	27.05 ± 0.08	29.41 ± 0.09	31.51 ± 0.12	31.60 ± 0.14
Ventral	30%	24.00 ± 0.12	33.21 ± 0.16	27.87 ± 0.09	25.51 ± 0.09	27.03 ± 0.05	29.37 ± 0.07	31.92 ± 0.11	32.06 ± 0.12
Lung	0%	23.69 ± 0.04	31.26 ± 0.14	30.17 ± 0.03	26.73 ± 0.06	26.82 ± 0.02	28.61 ± 0.07	30.77 ± 0.07	30.66 ± 0.09
Lung	30%	23.69 ± 0.09	28.81 ± 0.17	28.84 ± 0.07	26.71 ± 0.07	26.91 ± 0.04	28.34 ± 0.12	29.41 ± 0.12	29.16 ± 0.16

Supplementary Table 2.2 GeNorm candidate reference gene stability rankings for eight tissues from dehydrated *Xenopus laevis*. Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Stability (M) values are in parenthesis

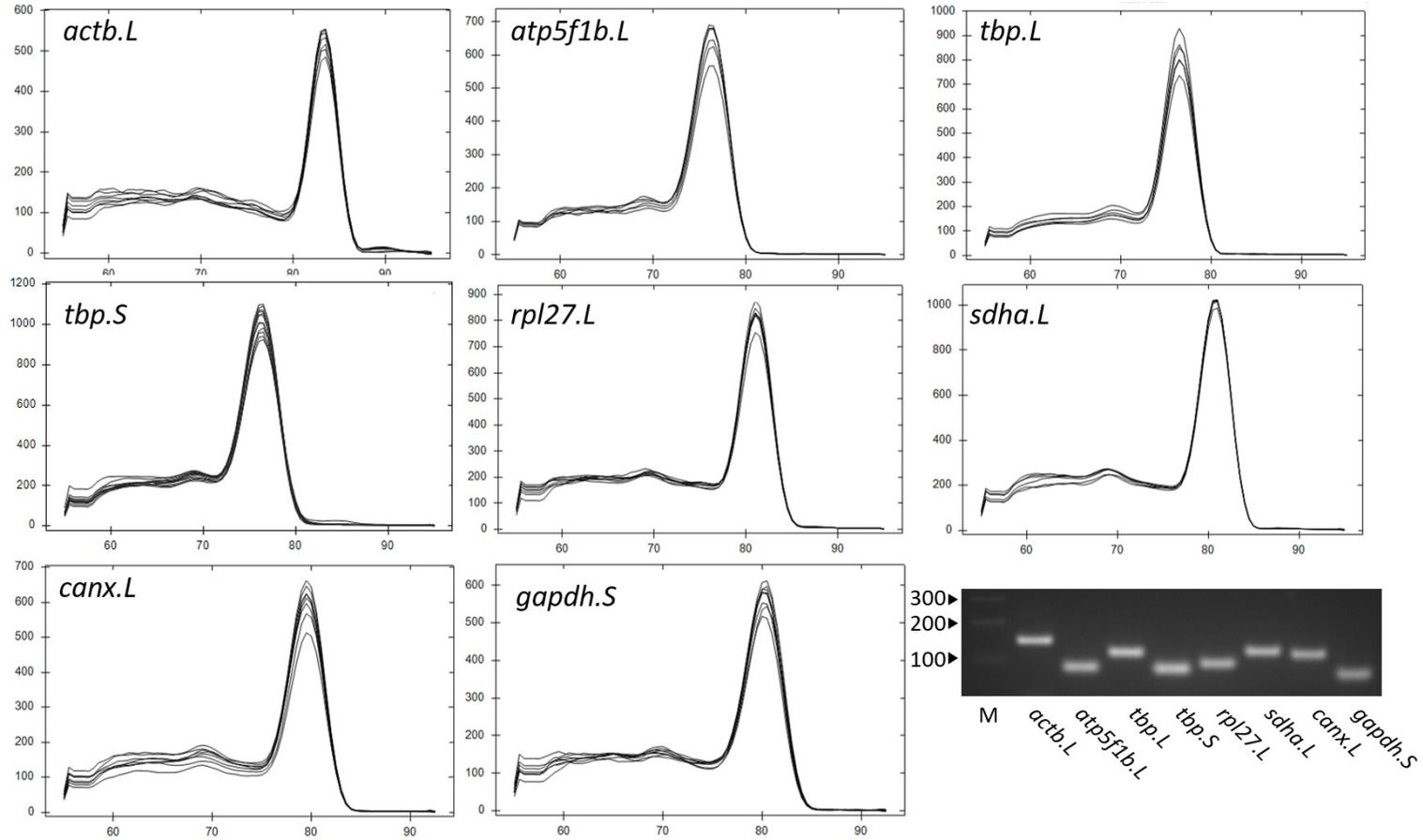
Rank	Brain	Dorsal	Heart	Kidney	Liver	Lung	Muscle	Ventral
1	<i>atp5f1b.L</i> (0.20)	<i>gapdh.S</i> (0.19)	<i>gapdh.S</i> (0.42)	<i>actb.L</i> (0.56)	<i>canx.L</i> (0.31)	<i>actb.L</i> (0.19)	<i>gapdh.S</i> (0.27)	<i>rpl27.L</i> (0.28)
1	<i>tbp.L</i> (0.20)	<i>sdha.L</i> (0.19)	<i>rpl27.L</i> (0.42)	<i>gapdh.S</i> (0.56)	<i>sdha.L</i> (0.31)	<i>gapdh.S</i> (0.19)	<i>rpl27.L</i> (0.27)	<i>tbp.L</i> (0.28)
3	<i>tbp.S</i> (0.23)	<i>rpl27.L</i> (0.31)	<i>actb.L</i> (0.51)	<i>tbp.L</i> (0.77)	<i>atp5f1b.L</i> (0.38)	<i>rpl27.L</i> (0.20)	<i>actb.L</i> (0.30)	<i>canx.L</i> (0.32)
4	<i>rpl27.L</i> (0.32)	<i>canx.L</i> (0.35)	<i>canx.L</i> (0.59)	<i>tbp.S</i> (0.91)	<i>actb.L</i> (0.47)	<i>sdha.L</i> (0.26)	<i>canx.L</i> (0.32)	<i>tbp.S</i> (0.36)
5	<i>canx.L</i> (0.35)	<i>actb.L</i> (0.41)	<i>tbp.L</i> (0.63)	<i>sdha.L</i> (1.00)	<i>gapdh.S</i> (0.52)	<i>canx.L</i> (0.47)	<i>tbp.S</i> (0.36)	<i>sdha.L</i> (0.42)
6	<i>gapdh.S</i> (0.40)	<i>atp5f1b.L</i> (0.57)	<i>tbp.S</i> (0.65)	<i>canx.L</i> (1.06)	<i>rpl27.L</i> (0.57)	<i>tbp.L</i> (0.56)	<i>tbp.L</i> (0.39)	<i>actb.L</i> (0.51)
7	<i>actb.L</i> (0.44)	<i>tbp.S</i> (0.70)	<i>atp5f1b.L</i> (0.68)	<i>atp5f1b.L</i> (1.12)	<i>tbp.L</i> (0.64)	<i>tbp.S</i> (0.63)	<i>sdha.L</i> (0.51)	<i>gapdh.S</i> (0.58)
8	<i>sdha.L</i> (0.52)	<i>tbp.L</i> (0.88)	<i>sdha.L</i> (0.74)	<i>rpl27.L</i> (1.25)	<i>tbp.S</i> (0.69)	<i>atp5f1b.L</i> (0.74)	<i>atp5f1b.L</i> (0.64)	<i>atp5f1b.L</i> (0.64)

Supplementary Table 2.3 Normfinder candidate reference gene stability rankings for eight tissues from dehydrated *Xenopus laevis*. Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Stability values are in parenthesis

Rank	Brain	Dorsal	Heart	Kidney	Liver	Lung	Muscle	Ventral
1	<i>canx.L</i> (0.06)	<i>rpl27.L</i> (0.11)	<i>actb.L</i> (0.21)	<i>actb.L</i> (0.2)	<i>atp5f1b.L</i> (0.18)	<i>canx.L</i> (0.3)	<i>rpl27.L</i> (0.19)	<i>canx.L</i> (0.14)
2	<i>rpl27.L</i> (0.07)	<i>sdha.L</i> (0.26)	<i>gapdh.S</i> (0.27)	<i>gapdh.S</i> (0.23)	<i>canx.L</i> (0.21)	<i>tbp.L</i> (0.38)	<i>gapdh.S</i> (0.21)	<i>rpl27.L</i> (0.19)
3	<i>gapdh.S</i> (0.11)	<i>canx.L</i> (0.27)	<i>rpl27.L</i> (0.29)	<i>tbp.L</i> (0.31)	<i>gapdh.S</i> (0.23)	<i>sdha.L</i> (0.52)	<i>canx.L</i> (0.24)	<i>tbp.L</i> (0.24)
4	<i>actb.L</i> (0.18)	<i>gapdh.S</i> (0.3)	<i>tbp.L</i> (0.31)	<i>atp5f1b.L</i> (0.48)	<i>sdha.L</i> (0.24)	<i>tbp.S</i> (0.54)	<i>tbp.L</i> (0.28)	<i>sdha.L</i> (0.26)
5	<i>tbp.S</i> (0.19)	<i>actb.L</i> (0.35)	<i>canx.L</i> (0.31)	<i>sdha.L</i> (0.51)	<i>actb.L</i> (0.26)	<i>rpl27.L</i> (0.56)	<i>tbp.S</i> (0.29)	<i>tbp.S</i> (0.36)
6	<i>atp5f1b.L</i> (0.2)	<i>tbp.S</i> (0.38)	<i>tbp.S</i> (0.34)	<i>tbp.S</i> (0.52)	<i>tbp.L</i> (0.33)	<i>actb.L</i> (0.56)	<i>actb.L</i> (0.34)	<i>atp5f1b.L</i> (0.39)
7	<i>tbp.L</i> (0.23)	<i>atp5f1b.L</i> (0.46)	<i>atp5f1b.L</i> (0.39)	<i>canx.L</i> (0.55)	<i>tbp.S</i> (0.34)	<i>gapdh.S</i> (0.58)	<i>sdha.L</i> (0.47)	<i>actb.L</i> (0.44)
8	<i>sdha.L</i> (0.34)	<i>tbp.L</i> (0.61)	<i>sdha.L</i> (0.45)	<i>rpl27.L</i> (0.66)	<i>rpl27.L</i> (0.36)	<i>atp5f1b.L</i> (0.99)	<i>atp5f1b.L</i> (0.54)	<i>gapdh.S</i> (0.5)

Supplementary Table 2.4 BestKeeper candidate reference gene stability rankings for eight tissues from dehydrated *Xenopus laevis*. Candidate reference genes are ranked from most stable (rank 1) to least stable (rank 8). Standard deviation values are in parenthesis

Rank	Brain	Dorsal	Heart	Kidney	Liver	Lung	Muscle	Ventral
1	<i>canx.L</i> (0.15)	<i>rpl27.L</i> (0.17)	<i>actb.L</i> (0.34)	<i>rpl27.L</i> (0.48)	<i>rpl27.L</i> (0.15)	<i>rpl27.L</i> (0.09)	<i>gapdh.S</i> (0.13)	<i>rpl27.L</i> (0.18)
2	<i>rpl27.L</i> (0.25)	<i>actb.L</i> (0.23)	<i>rpl27.L</i> (0.42)	<i>gapdh.S</i> (0.8)	<i>actb.L</i> (0.27)	<i>actb.L</i> (0.2)	<i>actb.L</i> (0.17)	<i>sdha.L</i> (0.24)
3	<i>actb.L</i> (0.33)	<i>canx.L</i> (0.26)	<i>atp5f1b.L</i> (0.43)	<i>actb.L</i> (1)	<i>sdha.L</i> (0.45)	<i>gapdh.S</i> (0.22)	<i>rpl27.L</i> (0.22)	<i>canx.L</i> (0.31)
4	<i>gapdh.S</i> (0.36)	<i>sdha.L</i> (0.3)	<i>gapdh.S</i> (0.44)	<i>tbp.L</i> (1.13)	<i>gapdh.S</i> (0.5)	<i>sdha.L</i> (0.32)	<i>canx.L</i> (0.25)	<i>tbp.L</i> (0.35)
5	<i>atp5f1b.L</i> (0.4)	<i>gapdh.S</i> (0.34)	<i>canx.L</i> (0.5)	<i>atp5f1b.L</i> (1.15)	<i>atp5f1b.L</i> (0.51)	<i>tbp.L</i> (0.64)	<i>tbp.S</i> (0.4)	<i>tbp.S</i> (0.46)
6	<i>tbp.L</i> (0.44)	<i>tbp.S</i> (0.57)	<i>tbp.L</i> (0.56)	<i>tbp.S</i> (1.34)	<i>canx.L</i> (0.55)	<i>canx.L</i> (0.65)	<i>tbp.L</i> (0.49)	<i>gapdh.S</i> (0.49)
7	<i>tbp.S</i> (0.45)	<i>atp5f1b.L</i> (0.69)	<i>sdha.L</i> (0.63)	<i>sdha.L</i> (1.37)	<i>tbp.L</i> (0.69)	<i>tbp.S</i> (0.73)	<i>sdha.L</i> (0.59)	<i>actb.L</i> (0.6)
8	<i>sdha.L</i> (0.53)	<i>tbp.L</i> (0.98)	<i>tbp.S</i> (0.69)	<i>canx.L</i> (1.41)	<i>tbp.S</i> (0.7)	<i>atp5f1b.L</i> (1.19)	<i>atp5f1b.L</i> (0.81)	<i>atp5f1b.L</i> (0.61)



Supplementary Figure 2.1 Reference gene qPCR primer validation. Melting curve derivative plots show a single prominent peak for each gene indicating a single product was amplified. Analysis using a 3% agarose gel (bottom right) show PCR products at expected sizes

Chapter 3

Changes to global DNA and histone modifications in response to dehydration in *Xenopus laevis*

Liam J. Hawkins¹ and Kenneth B. Storey¹

¹ Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

3.1 Abstract

The African clawed frog (*Xenopus laevis*) survives high levels of dehydration when its aquatic habitats dry up during arid summer months. Multiple adaptations contribute to dehydration tolerance in this species, including turning on the urea cycle, switching to anaerobic metabolism, expression of antioxidant proteins, among others. These adaptations require changes to gene expression that may be mediated by epigenetic mechanisms comprised of DNA and histone modifications that have been shown to be responsive in multiple other stress tolerant species, including other anurans. Here we measure genome-wide levels of epigenetic marks including 5mC, 5hmC, and histone modifications in the liver and skeletal muscle to determine if they may contribute to dehydration tolerance in *X. laevis*. Although no changes in 5mC were observed, increases in 5hmC in the liver may contribute to expression of liver-specific genes. Likewise, expression of *tet* genes, which code for DNA demethylating enzymes, increased by more than an order of magnitude in skeletal muscle, potentially due to their hypoxia inducibility. Lastly, we find little change to histone modifications in response to dehydration which contrasts other stress tolerant species. Together these results suggest potential fine-tuning of gene expression by epigenetic mechanisms and provides future directions for the study of dehydration tolerance in *X. laevis*.

3.2 Introduction

The African clawed frog, *Xenopus laevis*, is known for its uses as a model organism in the fields of developmental and cellular biology. It is however understudied regarding its adaptations to harsh environmental conditions, namely in the face of severe dehydration which this aquatic animal encounters seasonally. In its natural range of southern Africa, hot summer climates result in the evaporation of the lakes and ponds it resides in. As a result *X. laevis* demonstrate multiple behavioural and physiological adaptations to survive periods of severe dehydration. For example, when migration to a new water body is not possible, these animals burrow into the mud and secrete a layer of mucus to minimize evaporative water loss (Hillman 1978a). Regardless, the inevitable water loss results in decreased plasma volume and increased hematocrit, which is compensated by increased heart rate (Hillman 1978b). Measurements of lactate in multiple tissues indicate oxygen transport is impaired and that there is a shift from aerobic metabolism to anaerobic glycolysis as a source of ATP (Hillman 1978b). Also, nitrogen that is usually excreted in the form of ammonia is instead shunted into urea biosynthesis allowing it to accumulate to much higher levels until rehydration occurs (Balinsky et al. 1961; Hillman 1978a).

These physiological adaptations are accompanied by biochemical adaptations and changes in gene expression. For example, shifts to anaerobic glycolysis is supported by modifications to hexokinase, pyruvate kinase, and lactate dehydrogenase to facilitate glucose entry and flux in a tissue specific manner (Katzenback et al. 2014; Childers and Storey 2016; Dawson et al. 2018). The impaired oxygen transport may also result in increased oxidative damage, which is supported by the observed activation of antioxidant mechanisms (Malik and Storey 2009a, 2011), as well as increases in heat shock proteins

(Luu et al. 2018). Recently, microRNA, small non-coding RNA molecules which can specifically target and silence mRNA from being transcribed have also been shown to be dehydration responsive, and potentially contribute to a neuroprotective phenotype (Luu and Storey 2015).

Adaptations involving changes to gene expression require mechanisms to coordinate these responses. Epigenetic mechanisms such as DNA methylation and histone modifications provide global control of gene expression across the genome, in contrast to specialized trans-regulatory elements such as transcription factors with specific binding motifs. DNA methylation involves the transfer of a methyl group to the fifth position of cytosine nucleotides (5-methylcytosine; 5mC), typically in the CpG context. This transfer is catalyzed by DNA methyltransferases (DNMTs). DNMTs are considered to be either maintenance methyltransferases, or *de novo* methyltransferases, although this distinction is starting to be questioned (Jeltsch and Jurkowska 2014). DNMT1 is the primary maintenance methyltransferase since it copies existing methylation patterns onto newly synthesized DNA strands, whereas DNMT3A and DNMT3B are *de novo* methyltransferases – integrating signals from the environment into new DNA methylation patterns. DNMT3L, while having no enzymatic activity, aids the other DNMTs in targeting and activity of their methyltransferase activity (Kareta et al. 2006).

DNA methylation patterns are dynamic, and the removal of 5mC begins with its oxidation to form 5-hydroxymethylcytosine (5hmC) by the ten-eleven translocation (TET) class of enzymes. Once converted from 5mC to 5hmC, it is no longer recognized by DNMT1 and can then be passively demethylated by dilution over the course of multiple cell cycle rounds. Alternatively, 5hmC can continue to be serially oxidized by

TET enzymes to form 5-formylcytosine and 5-carboxycytosine where they can then be excised by the DNA repair enzyme thymine DNA glycosylase (TDG) and repaired to an unmodified cytosine (He et al. 2011; Ito et al. 2011; Maiti and Drohat 2011).

When present at gene promoters and transcription start sites, DNA methylation is associated with transcriptional silencing. The recruitment of methyl-binding domain containing proteins sterically hinder transcription factor binding thereby inhibiting their induction. However, it has also been observed that gene bodies of actively transcribed genes are highly methylated, and this methylation stabilizes the transcriptional machinery and prevents spurious transcription. While the role of DNA methylation is dependent on location within a gene, hypomethylation phenotypes are associated with global increases in transcription and vice-versa (Hawkins et al. 2018a), suggesting that global measurements of DNA methylation levels can indicate global transcriptional status, even when considering the position dependant function of DNA methylation.

Like DNA, histones can be covalently modified to affect gene transcription. Their effect on transcription is dependent on: (1) the type of modification, (2) which histone subunit is modified, (3) the location on the histone subunit, and (4) the local histone and DNA modification context (Bannister and Kouzarides 2011). The most well characterized modifications are histone acetylation and methylation. In the case of acetylation, acetyl groups donated from acetyl-CoA are transferred onto lysine residues removing their positive charge. This change in charge weakens the histone-DNA interaction and promotes the loosening of chromosomal structures and increased DNA accessibility (Lee et al. 1993). This, combined with the recruitment of transcriptional

machinery typically associates histone acetylation with transcriptional activation (Lawrence et al. 2016).

Like DNA methylation, S-adenosylmethionine is the methyl-donor for histone methylation, and up to three methyl groups can be added to each lysine residue. Unlike acetylation, histone methylation does not change the histone-DNA electrostatic interaction and therefore is correlated with a range of transcriptional outcomes depending on the degree of methylation and location on a histone subunit (Kouzarides 2007). For example, the methylation of histone H3 lysine 4 (H3K4) is typically associated with actively transcribed genes, whereas tri-methylated histone H3 lysine 9 (H3K9me3) is correlated with transcriptional silencing (Kouzarides 2007; Lawrence et al. 2016).

Whether epigenetic mechanisms like histone modifications and DNA methylation contribute to dehydration tolerance is unknown. Results from other animal species adapted to myriad environmental stressors show changes to epigenetic marks and machinery (Hawkins and Storey 2020), therefore, it appears probable that these same mechanisms may be involved in dehydration tolerance in *X. laevis*. Since *X. laevis* show changes to gene expression in response to dehydration, this study seeks to determine whether global levels of epigenetic marks, namely DNA methylation and histone modifications, are affected during dehydration.

3.3 Materials and Methods

3.3.1 Animal treatment

Adult male *X. laevis* from the University of Alberta Science Animal Support Services were acclimated in tanks of dechloraminated water at 22°C for one week. Animals were fed three times during this period and the water was changed after feeding. Subsequently,

animals were divided into two groups, control and dehydration. Control animals were kept in the above conditions for the remainder of the experiment. Animals in the dehydration group were transferred to dry containers at 22°C and were periodically weighed over several days until they had lost approximately 30% of their body water according to the following equation:

$$\% \text{ Body water loss} = \frac{m_i - m_d}{m_i \times BWC_i} \times 100\%$$

where m_i is the initial mass of the animal, m_d is the dehydrated mass, and BWC_i is initial body water content as previously determined (Malik and Storey 2009b). Animals from both groups were not fed for the duration of the dehydration period and were then euthanized by double pithing and tissues were quickly dissected, flash frozen in liquid nitrogen, and stored at -80°C until use. All protocols were approved by the Carleton University Animal Care Committee (protocol #106936) and conformed within the guidelines of the Canadian Council on Animal Care.

3.3.2 Genomic DNA extraction

Genomic DNA was extracted from 25 mg of frozen tissue (n=6 per experimental condition per tissue) using the Quick-DNA Miniprep Plus Kit (Zymo Research Cat#: D4068) according to the manufacturer's directions for tissue samples. Briefly, tissues were digested with proteinase K overnight at 55°C, then centrifuged at 12,000 x g for 1 min and aqueous supernatant containing DNA was transferred and bound to the supplied spin column. Spin columns were washed multiple times with the provided wash buffer, then eluted in the supplied DNA elution buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.5). Genomic DNA samples were then measured using a Take3 Micro-Volume plate

(BioTek, Vermont, USA) and all samples had a 260/280 ratio > 1.8. DNA integrity was validated on a 0.5% agarose gel.

3.3.3 Quantification of DNA modifications

Global 5mC levels were measured using the 5mC DNA ELISA Kit (Zymo Research Cat#: D5325) according to the manufacturer's directions. Briefly, DNA samples were denatured at 98°C then cooled on ice before binding samples to wells of the supplied 96-well plate. Each well was then blocked using the supplied 5mC ELISA blocking buffer and a mixture of anti-5mC primary antibodies and secondary detection antibodies was added to each well and incubated at 37°C for one hour. The supplied HRP Developer was then added to the wells and absorbance was read at 450 nm. A standard curve of known 5mC quantities supplied by the manufacturer was run in parallel and 5mC percentages for each sample were calculated using the equation from a logarithmic regression of the standard curve.

Global 5hmC levels were measured using the Global 5hmC Quantification Kit (Active Motif Cat#: 55018) according to the manufacturer's directions. Briefly, 0.5 µg of each DNA sample was digested using the supplied MseI restriction digest enzyme at 37°C for 3 h. Digested DNA was then quantified and normalized using the Take3 Micro-Volume plate. Digested DNA samples and supplied 5hmC standards were then denatured at 98°C for 10 min and chilled on ice before binding to the wells of the supplied 96-well plate. Each well was washed three times with the supplied wash buffer and anti-5hmC primary antibody was added to each well for 1 h at room temperature. Wells were then washed three more times and the HRP-conjugated secondary antibody was added to each well for 1 h at room temperature. The colorimetric reaction was performed by adding the

supplied developing solution for 3 min, then stopped using the supplied stop solution and read at 450 nm. The 5hmC percentage of each sample was then calculated using the standard curve.

3.3.4 Nuclear fractionation and protein extraction

Liver and skeletal muscle tissue samples (n=4 per experimental condition per tissue) were weighed then crushed in liquid N₂ and homogenized with a Dounce homogenizer 1:5 w/v in chilled cytoplasmic fractionation buffer (10 mM HEPES, 10 mM KCl, 10 mM EDTA, 20 mM β-glycerol phosphate, 1 mM DTT, 0.1% Triton X-100, and 10 μL/mL of protease inhibitor cocktail, pH 7.9), then vortexed briefly and incubated on ice for 60 min. Samples were then centrifuged at 10,000 x g for 15 min at 4°C and the supernatant containing the cytoplasmic fraction was removed. The pellet was then washed three times with cytoplasmic buffer and resuspended in nuclear fractionation buffer (20 mM HEPES, 400 mM NaCl, 1 mM EDTA, 2% v:v glycerol, 20 mM β-glycerol phosphate, pH 7.9). The sample was then sonicated and briefly vortexed, then incubated on ice for 15 min and centrifuged at 10,000 x g for 15 min at 4°C. The supernatant containing nuclear proteins was removed and nuclear and cytoplasmic fractions were normalized to a common concentration. Samples were then stored at -80°C until use in enzymatic assays, or mixed 1:1 v/v with SDS buffer (100 mM Tris-base, 4% w/v SDS, 20% v/v glycerol, 0.2% w/v bromophenol blue, 10% v/v 2-mercaptoethanol) and boiled in a water bath for 10 min to fully denature and linearize proteins in preparation for western blotting. Samples were then probed with a cytoplasmic (β-tubulin) and nuclear (H3K4me1) specific antibodies to validate fractionation.

3.3.5 DNMT activity assay

DNMT activity was quantified using the EpiQuik DNMT Activity/Inhibition Assay Ultra Kit (Epigentek Cat#: P-3009-96). Briefly, protein samples were added to wells of the supplied 96-well plate coated with DNMT substrates and incubated for 2 h at 37°C. Wells were then washed three times with the supplied wash buffer, before the supplied anti-5mC primary antibody was added to each well and incubated for 1 h at room temperature. The wells were washed three more times and the supplied secondary detection antibody was added to each well and incubated for 30 min at room temperature. Wells were then washed four more times and the supplied enhancer solution was incubated for 30 min in each well at room temperature before washing five times. The supplied developer solution was incubated in each well until the positive control turned dark blue at which point the supplied Stop Solution was added to each well and absorbance was read at 450 nm.

3.3.6 RNA extraction and cDNA synthesis

RNA extraction and cDNA synthesis was performed as previously described (Hawkins et al. 2018b). Briefly, 50 mg of tissue samples were homogenized in TRI Reagent (Molecular Research Center), then a chloroform-isopropanol RNA extraction was performed. Reverse transcription of extracted RNA was performed using the M-MLV Reverse Transcriptase Kit (Cat.# 28025013) and samples were stored at -20°C until use.

3.3.7 Primer design and RT-qPCR

Primers were designed using Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) to generate a PCR product size of 80-

120bp, unless otherwise specified, and checked for *Xenopus laevis* primer pair specificity (taxid: 8355). Primers were synthesized by Integrated DNA Technologies and are available in **Table 3.1**.

RT-qPCR was performed using a Bio-Rad CFX384 Touch as previously described (Pellissier et al. 2006). PCR reactions contained 4 μL of cDNA, 2 μL qPCR buffer (100 mM Tris-HCl [pH 8.5], 500 mM KCl, 1.5% v:v Triton X-100, and 20 mM MgCl_2), 0.08 μL of 25 mM dNTPs, 2 μL of 1 M trehalose, 0.25 μL of formamide, 0.0125 μL of 100X SYBR Green in DMSO (Invitrogen; #S7585), 0.25 μL of 0.3 $\text{nmol}\cdot\mu\text{L}^{-1}$ of each primer, 0.0625 μL of 5U $\cdot\mu\text{L}^{-1}$ wild-type Taq (BioShop; #TAQ001.1) and autoclaved ddH₂O for a final volume of 10 μL . The following thermocycle series was used: An initial 95°C denaturing step for 3 min, then 40 cycles of denaturing at 95°C for 10 sec, annealing at 57°C for 20 sec, and extension at 72°C for 20 sec. A melting analysis was performed from 55-95°C to ensure the amplification of single products, and a 2-fold dilution standard curve for each target was run on quantification plates to ensure reaction efficiencies between 90-110%. Suitable reference genes were previously determined to be *atp5f1b.L* and *sdha.L* for liver and *gapdh.S* and *rpl27.L* for skeletal muscle (Hawkins et al. 2018b). Quantification was performed using Bio-Rad CFX Maestro 1.1 which uses the Pfaffl method (Pfaffl 2001).

3.3.8 Western blotting

Western blotting was performed as previously described (Hawkins and Storey 2018). Briefly, nuclear protein extracts were loaded into 10-15% SDS-PAGE gels and run for 60-120 min at 180 V depending on the size of the protein to be probed with a Mini Protean III system (BioRad Laboratories, Hercules, CA, USA). Proteins were then

transferred onto PVDF membranes using a wet-transfer apparatus (BioRad Laboratories, Hercules, CA, USA). Membranes were then washed with TBST, blocked with skim milk in TBST for 30 min, washed three more times and probed with primary antibodies overnight at 4°C. Membranes were then washed three times, probed with secondary antibody for 30 min at room temperature, washed three more times and visualized with chemiluminescence using the ChemiGenius Bio Imaging System (Syngene, Frederick, MD, USA). Membranes were then stained with Coomassie blue for protein loading control (Welinder and Ekblad 2011).

3.3.9 Statistics

Statistical significance testing was performed using Student's t-test to compare samples from control and dehydrated animals using SigmaPlot 12 statistical package software (Systat Software Inc., San Jose, CA, USA). A *p*-value < 0.05 was considered significant. Data are expressed as mean ± SEM.

3.4 Results

3.4.1 Global DNA modification levels

Changes in 5mC between control and dehydrated animals were measured in liver and skeletal muscle and remained unchanged during dehydration (**Figure 3.1a**). Both tissues showed 5mC levels typical of vertebrates (3.28 ± 0.66 and 4.26 ± 0.38 % 5mC in liver and skeletal muscle of control animals respectively). In contrast, liver 5hmC levels showed a statistically significant difference during dehydration (**Figure 3.1b**). In this case, liver 5hmC levels increased from 0.38 ± 0.04 % 5hmC in control animals to 0.87 ± 0.07 in dehydrated animals. The levels of 5hmC measured here are also consistent with

levels seen in other vertebrate tissues (Dahl et al. 2011). Muscle 5hmC levels were not significantly affected by dehydration.

3.4.2 DNMT activity levels

DNMT activity levels were measured in liver and skeletal muscle comparing control to dehydrated animals (**Figure 3.2**). Tissue level differences in DNMT activity from control animals showed that liver had approximately 6-fold higher activity levels than skeletal muscle. However, skeletal muscle was the only tissue that showed statistically significant differences between experimental conditions, where DNMT activities in dehydrated animals increased 3.59 ± 0.87 -fold over control animals.

3.4.3 *Dnmt* expression

Although no changes in global 5mC or DNMT activity levels were observed, expression levels of *dnmt* genes were significantly downregulated in the liver during dehydration (**Figure 3.3a**). Specifically, both long chromosome (L) DNMT genes, *dnmt1.L* and *dnmt3a.L*, decreased to 0.32 ± 0.11 and 0.56 ± 0.12 of controls levels respectively, whereas the short chromosome (S) versions were unchanged. None of the *dnmt1* or *dnmt3a* genes showed a significant response in skeletal muscle to dehydration (**Figure 3.3b**). It should also be noted that no *dnmt3b* or *dnmt3l* genes are found in the *X. laevis* genome.

3.4.4 DNA demethylation gene expression

Genes involved in DNA demethylation showed significant regulation in response to dehydration in both tissues (**Figure 3.4**). In the liver, only *tet3.S* was significantly downregulated during dehydration from 1.26 ± 0.61 to 0.48 ± 0.08 , while in skeletal

muscle *tet2.L* increased from 1.00 ± 0.37 to 14.43 ± 2.25 of control, *tet2.S* increased from 1.45 ± 0.50 to 15.32 ± 2.10 , and *tet3.L* increased from 1.00 ± 0.14 to 1.69 ± 0.18 . The gene for TDG only exists as a short chromosome form (*tdg.S*) and was unchanged in both tissues during dehydration. The gene for *tet1* is not known to exist in *X. laevis*.

3.4.5 Histone modifications

We measured relative levels of histone modifications (methylation and acetylation) using western blotting and in each case the levels of modification were relative to levels of total histone H3 levels also determined using western blotting. In the liver (**Figure 3.5**) the only modification that significantly changed during dehydration was H3K4me, a mark that is typically associated with active transcription. No measured histone modification showed significant changes in skeletal muscle during dehydration (**Figure 3.6**).

3.5 Discussion

X. laevis naturally survive severe dehydration as their aquatic habitats evaporate during summer months. Their ability to do so is predicated on multiple tissue level responses including a transition from aerobic to anaerobic metabolism as oxygen transport decreases with increased blood viscosity and an upregulation of the urea cycle as the animal switches from ammonotelism to ureotelism. These adaptations and others are, unsurprisingly, in part a result of changes in gene expression. For example, pyruvate kinase, an enzyme whose regulation is typically studied at the allosteric and covalent level, is upregulated at the mRNA level in liver and skeletal muscle during dehydration (Dawson et al. 2018). Likewise, transcripts and protein levels of antioxidant enzymes are also upregulated in the liver (Malik and Storey 2011). DNA and histone modifications are

mechanisms for transducing intra- and extracellular signals into transcriptional programming, and therefore are candidates for regulation of gene expression in response to dehydration in *X. laevis*.

3.5.1 DNA methylation status during dehydration

We started by measuring global levels of the two most common DNA modifications, 5mC and 5hmC, the latter being an indication of DNA demethylation as it is the first step in the serial oxidation and eventual demethylation of cytosines (Wu and Zhang 2017). While 5mC levels were not altered in either tissue during dehydration, 5hmC levels in the liver almost doubled (**Figure 3.1**). Increased 5hmC levels without changes in 5mC levels may indicate a rearrangement of methylation pattern in liver. Given that 5mC levels were measured at nearly an order of magnitude greater than 5hmC, any reduction in 5mC as a result of oxidation may not reach measurable levels in this case where 5hmC levels increase by approximately 0.5%. If DNA methylation is altered to control gene expression in response to dehydration, then the removal of 5mC from some genes and the addition to others may not result in large changes to 5mC levels, but would temporarily result in a rise in 5hmC levels relative to the low basal levels of 5hmC (Dahl et al. 2011).

The 5hmC modification is also known to be present on actively expressed tissue-specific genes, but not highly expressed “housekeeping” genes (Pfeifer and Szabó 2018). In the context of our results, this may indicate that the increase in liver 5hmC levels is correlated to liver-specific genes being expressed at higher levels during dehydration. This is consistent with the known physiological and metabolic response of the liver during dehydration which includes turning on the urea cycle (Balinsky et al. 1961) and altering amino acid metabolism (Balinsky et al. 1967), both highly-specific to the liver.

The lack of change in 5hmC levels in the skeletal muscle is also consistent here: although many metabolic responses have been identified in skeletal muscle, no skeletal muscle-specific processes are known to occur in response to dehydration in *X. laevis*. During dehydration, the animals assume a water-conserving posture and show little movement, likely to avoid using energy and exposing additional surface area, therefore for the duration of dehydration there is little use for skeletal muscle activity which may explain why more tissue-specific metabolic responses have been identified in liver as compared to muscle.

These results contrast what is seen in another stress tolerant anuran species, the wood frog (*Rana sylvatica*) (Zhang et al. 2019). This species tolerates whole body freezing each winter, and we previously showed that 5mC levels increase while 5hmC levels decrease in skeletal muscle during freezing. Likewise, the anoxia tolerant red eared slider (*Trachemys scripta elegans*) also showed increased 5mC in both the liver and white muscle when exposed to 5 and 20 h of anoxia (Wijenayake and Storey 2016). In both cases, these species show metabolic rate depression which necessitates a strong suppression of transcription and translation to survive their respective stresses (Storey and Storey 2004). The absence of a global metabolic suppression in dehydrated *X. laevis* may account for the lack of large genome wide changes in 5mC levels.

Interestingly DNMT activity showed an increase in skeletal muscle during dehydration (**Figure 3.2**), although this did not result in increased 5mC levels (**Figure 3.1**). This measured activity, of course, does not account for the highly regulated localization of activity of DNMT enzymes *in vivo*. It therefore may be the case that

increased activity is directed to specific genes and not the genome as a whole, although further analysis is needed to determine this.

3.5.2 Hypoxia-inducible TETs highly upregulated in skeletal muscle

Expression of liver *dnmt* genes did not correlate with 5mC levels or DNMT activity, although this is unsurprising given the number of regulatory steps between mRNA and protein functionality. Although no changes to 5mC and DNMT activity were seen in this tissue, mRNA of *dnmt1.L* and *dnmt3a.L* dropped to approximately half of control levels during dehydration (**Figure 3.3a**). A possible explanation for this is that under the stressful conditions of dehydration, the cell cycle may be reduced as seen in the dehydration tolerant wood frog (Roufayel et al. 2011). A reduction in cell division and DNA synthesis would require less DNMT1 which may explain why *dnmt1.L* levels decrease during dehydration. In contrast, expression of *dnmt* genes in skeletal muscle corresponded to 5mC levels in that neither changed significantly in response to dehydration (**Figure 3.1b; Figure 3.3b**), whereas neither correlated with the increased *in vitro* DNMT activity levels (**Figure 3.2**). Increased mRNA levels are just one contributing factor to the measured activity of an enzyme, and therefore various other explanations exist with regard to the divergence in results between mRNA levels and DNMT activities. For example, we previously showed that DNMT activity was uniquely regulated by glucose in the freeze tolerant wood frog (Zhang et al. 2019). Unique regulation by glucose in that species is eco-physiologically relevant due to its production of massive quantities of glucose to use as a cryoprotectant. Likewise, changes in the cellular metabolite milieu are bound to occur during dehydration in *X. laevis* due to

changes in solute concentrations, which opens the possibility of allosteric regulation of DNMTs by the microenvironment.

During dehydration, liver and muscle showed very different *tet* enzyme DNA demethylation profiles. *Tet* genes in liver were mostly unchanged except for *tet3.S* (**Figure 3.4a**), while most notably in skeletal muscle both sub-genome versions of *tet2* increased more than an order of magnitude (**Figure 3.4b**). This increase in *tet* expression could be explained by the known HIF-mediated upregulation of *tet* genes (Wu and Zhang 2017). During dehydration, whole body lactate levels increase, and oxygen delivery becomes the limiting factor for survivability (Hillman 1978b). This dehydration-induced hypoxia may also increase HIF1 α levels. With oxygen as a substrate, the HIF-mediated expression of TET enzymes allows for demethylation to occur in the face of decreased oxygen availability (Alivand et al. 2017). This mechanism of regulation appears to be tissue specific (Thienpont et al. 2016) which may explain the differences seen here in liver and skeletal muscle. Levels of 5hmC in skeletal muscle showed no statistical difference, but appear to be upregulated (**Figure 3.1b**). If this is in fact the case, it may be that increased *tet* expression in skeletal muscle would lead to increased 5hmC levels after a longer period of dehydration, although this would need to be determined experimentally.

3.5.3 Histone methylation and acetylation largely unaffected by dehydration

As with DNA modifications, histone modifications control expression of genes. The transcriptional effect depends on the type, placement, and local context of modifications resulting in combinatorial complexity that is known as the histone code (Allis and Jenuwein 2016). Although the exact function of a given mark is difficult to determine,

general correlated functions of histone modifications have been shown. For example, histone acetylation is typical of actively transcribed genes (Bannister and Kouzarides 2011), whereas histone methylation can be both a positive or negative regulator (Barski et al. 2007).

Here we measured multiple histone acetylation and methylation modifications and found little change in global levels between control and dehydrated animals (**Figure 3.5; Figure 3.6**). The only modification that showed a significant difference was mono-methylated histone H3 lysine 4 (H3K4me) in the liver (**Figure 3.5**), whereas no modifications significantly changed in skeletal muscle. H3K4me is typically associated with active transcription and as we have previously shown, another anuran species adapted to environmental stress also shows decreased H3K4me levels (Hawkins and Storey 2018). Unlike *Xenopus*, wood frog showed a more profound response in levels of other histone modifications, possibly due to the inherent differences in transcriptional status these two species assume while tolerating their respective stresses. Wood frogs undergo significant metabolic rate reduction, whereas *X. laevis* do not reduce metabolic rate to survive dehydration and therefore global suppression of gene expression is unlikely. This may explain the discordance between these two animals, particularly in skeletal muscle tissue where *X. laevis* displayed no change in histone modifications, whereas levels of the transcriptional activating H3K4me and H3K27me were almost undetectable in metabolically depressed wood frogs during freezing exposure (Hawkins and Storey 2018). Although large fluctuations in global levels were minimal, this does not rule out potential important changes to local epigenetic mark levels at individual genes.

3.5.4 Conclusions

Our genome-wide examination of important epigenetic markers in *X. laevis* during dehydration revealed tissue-specific responses in liver and muscle. Although no changes in 5mC was seen in either tissue, increased levels of 5hmC in the liver may contribute to the expression of tissue-specific genes known to occur during dehydration. We also showed that genes for TET enzymes increased expression by more than an order of magnitude in skeletal muscle which may be due to hypoxia inducible regulation of these genes. Finally, we showed very little histone modifications changes occur during dehydration in contrast to other stress tolerant anurans. Together these results provide a foundation for future studies to explore tissue- and gene-specific regulation of dehydration tolerance by epigenetic mechanisms in *X. laevis*.

3.6 References

- Alivand MR, Soheili Z-S, Pornour M, et al (2017) Novel Epigenetic Controlling of Hypoxia Pathway Related to Overexpression and Promoter Hypomethylation of TET1 and TET2 in RPE Cells. *J Cell Biochem* 118:3193–3204 . doi: 10.1002/jcb.25965
- Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. *Nat Rev Genet* 17:487–500 . doi: 10.1038/nrg.2016.59
- Balinsky JB, Choritz EL, Coe CG, van der Schans GS (1967) Amino acid metabolism and urea synthesis in naturally aestivating *Xenopus laevis*. *Comp Biochem Physiol* 22:59–68 . doi: 10.1016/0010-406x(67)90166-1
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Cell Res* 21:381–95 . doi: 10.1038/cr.2011.22
- Barski A, Cuddapah S, Cui K, et al (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837 . doi: 10.1016/j.cell.2007.05.009
- Childers CL, Storey KB (2016) Post-translational Regulation of Hexokinase Function and Protein Stability in the Aestivating Frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Dahl C, Grønbaek K, Guldberg P (2011) Advances in DNA methylation: 5-hydroxymethylcytosine revisited. *Clin Chim Acta* 412:831–6 . doi: 10.1016/j.cca.2011.02.013
- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- Hawkins LJ, Al-Attar R, Storey KB (2018a) Transcriptional regulation of metabolism in disease: From transcription factors to epigenetics. *PeerJ* 6:e5062 . doi: 10.7717/peerj.5062
- Hawkins LJ, Luu BE, Storey KB (2018b) Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. *Gene Reports* 13:192–198 . doi: 10.1016/j.genrep.2018.10.006
- Hawkins LJ, Storey KB (2020) Advances and applications of environmental stress adaptation research. *Comp Biochem Physiol A Mol Integr Physiol* 240:110623 . doi: 10.1016/j.cbpa.2019.110623
- Hawkins LJ, Storey KB (2018) Histone methylation in the freeze-tolerant wood frog (*Rana sylvatica*). *J Comp Physiol B* 188:113–125 . doi: 10.1007/s00360-017-1112-7

- He Y-F, Li B-Z, Li Z, et al (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333:1303–7 . doi: 10.1126/science.1210944
- Hillman SS (1978a) Some effects of dehydration on internal distributions of water and solutes in *Xenopus laevis*. *Comp Biochem Physiol Part A Physiol* 61:303–307
- Hillman SS (1978b) The roles of oxygen delivery and electrolyte levels in the dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Ito S, Shen L, Dai Q, et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333:1300–3 . doi: 10.1126/science.1210597
- Jeltsch A, Jurkowska RZ (2014) New concepts in DNA methylation. *Trends Biochem Sci* 39:310–318 . doi: 10.1016/j.tibs.2014.05.002
- Kareta MS, Botello ZM, Ennis JJ, et al (2006) Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *J Biol Chem* 281:25893–902 . doi: 10.1074/jbc.M603140200
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128:693–705
- Lawrence M, Daujat S, Schneider R (2016) Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* 32:42–56 . doi: 10.1016/j.tig.2015.10.007
- Lee DY, Hayes JJ, Pruss D, Wolffe AP (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72:73–84 . doi: 10.1016/0092-8674(93)90051-Q
- Luu BE, Storey KB (2015) Dehydration triggers differential microRNA expression in *Xenopus laevis* brain. *Gene* 573:64–9
- Luu BE, Wijenayake S, Malik AI, Storey KB (2018) The regulation of heat shock proteins in response to dehydration in *Xenopus laevis*. *Cell Stress Chaperones* 23:45–53 . doi: 10.1007/s12192-017-0822-9
- Maiti A, Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* 286:35334–8 . doi: 10.1074/jbc.C111.284620
- Malik AI, Storey KB (2011) Transcriptional regulation of antioxidant enzymes by FoxO1 under dehydration stress. *Gene* 485:114–9 . doi: 10.1016/j.gene.2011.06.014
- Malik AI, Storey KB (2009a) Activation of antioxidant defense during dehydration stress in the African clawed frog. *Gene* 442:99–107 . doi: 10.1016/j.gene.2009.04.007

- Malik AI, Storey KB (2009b) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Pellissier F, Glogowski CM, Heinemann SF, et al (2006) Lab assembly of a low-cost, robust SYBR green buffer system for quantitative real-time polymerase chain reaction. *Anal Biochem* 350:310–2 . doi: 10.1016/j.ab.2005.12.002
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45 . doi: 10.1093/nar/29.9.e45
- Pfeifer GP, Szabó PE (2018) Gene body profiles of 5-hydroxymethylcytosine: potential origin, function and use as a cancer biomarker. *Epigenomics* 10:1029–1032 . doi: 10.2217/epi-2018-0066
- Roufayel R, Biggar KK, Storey KB (2011) Regulation of cell cycle components during exposure to anoxia or dehydration stress in the wood frog, *Rana sylvatica*. *J Exp Zool A Ecol Genet Physiol* 315:487–94 . doi: 10.1002/jez.696
- Storey KB, Storey JM (2004) Metabolic rate depression in animals: Transcriptional and translational controls. *Biol Rev Camb Philos Soc* 79:207–233
- Thienpont B, Steinbacher J, Zhao H, et al (2016) Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* 537:63–68 . doi: 10.1038/nature19081
- Welinder C, Ekblad L (2011) Coomassie staining as loading control in Western blot analysis. *J Proteome Res* 10:1416–9 . doi: 10.1021/pr1011476
- Wijenayake S, Storey KB (2016) The role of DNA methylation during anoxia tolerance in a freshwater turtle (*Trachemys scripta elegans*). *J Comp Physiol B* 186:333–342 . doi: 10.1007/s00360-016-0960-x
- Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 18:517–534 . doi: 10.1038/nrg.2017.33
- Zhang J, Hawkins LJ, Storey KB (2019) DNA methylation and regulation of DNA methyltransferases in a freeze tolerant vertebrate. *Biochem Cell Biol*. doi: 10.1139/bcb-2019-0091

Table 3.1. Primers used for RT-qPCR analysis of *Xenopus laevis*. Both short (S) and long (L) chromosome sub-genome genes of each target were measured unless one version did not exist. NCBI RefSeq or nucleotide accession numbers are provided for each gene. Reference genes for normalization (*atp5f1b.L* and *sdha.L* for liver, *rpl27.L* and *gapdh.S* for skeletal muscle) were used as previously determined (Hawkins et al. 2018b).

Gene symbol	Accession	Forward Primer	Reverse Primer
<i>dnmt1.L</i>	NM_001090552	CTGCGCAACAGCACTTCTAA	AGTGGAGAGGTAAGACGCCA
<i>dnmt1.S</i>	NM_001123360	TGAATCATTTGTCGCTTCTGGT	CGCACAGCTCATGTGTTTCA
<i>dnmt3a.L</i>	XM_018262178	GCTAGCCTTGCCAGTCAGTAA	CCAAACGCCTTCTGGACTCA
<i>dnmt3a.S</i>	XM_018266063	GGCACCGGGAGACTCTTTTT	GGCCGTTTGTCTCCTTCCTT
<i>tet2.L</i>	XM_018251892	TTTAGCAGCAGCTATGGATTTATCA	TGGTTGGTTCTATCCTCTTCCA
<i>tet2.S</i>	XM_018243385	CCCTTCTGCTCTGAGGGAAC	TGTATACGTATCTGCAGAAGCAC
<i>tet3.L</i>	HQ220208	CGGAGTTGAATGGTGGTGGT	TGAGTTGCTCGCATTTTCGC
<i>tet3.S</i>	HQ220207	TCAGCCCACATCACTTCCAC	CTCCCAAGGACTTTCGGTC
<i>tdg.S</i>	NM_001090821	CCCAAACGCATGTTTTGAACC	ATCCAAACAAAGGCTCCACCT
<i>atp5f1b.L</i>	XM_018247436	TGCCCCTTTGTCTTGGGTAA	AACAGCCACACAAGTTGGTC
<i>sdha.L</i>	NM_001090004	TCACGGCTTCTTAGCAGAGC	TGCTGAAGTGTCTTCCGTC
<i>rpl27.L</i>	NM_001086954	GAGTTGGAGTGAAAATGGGCA	TCACAATGACAGCTTTCGGG
<i>gapdh.S</i>	NM_001087098	GCCGCCATTAAGACTGCATCA	TGGAGACAACCTGGTCTTGTG

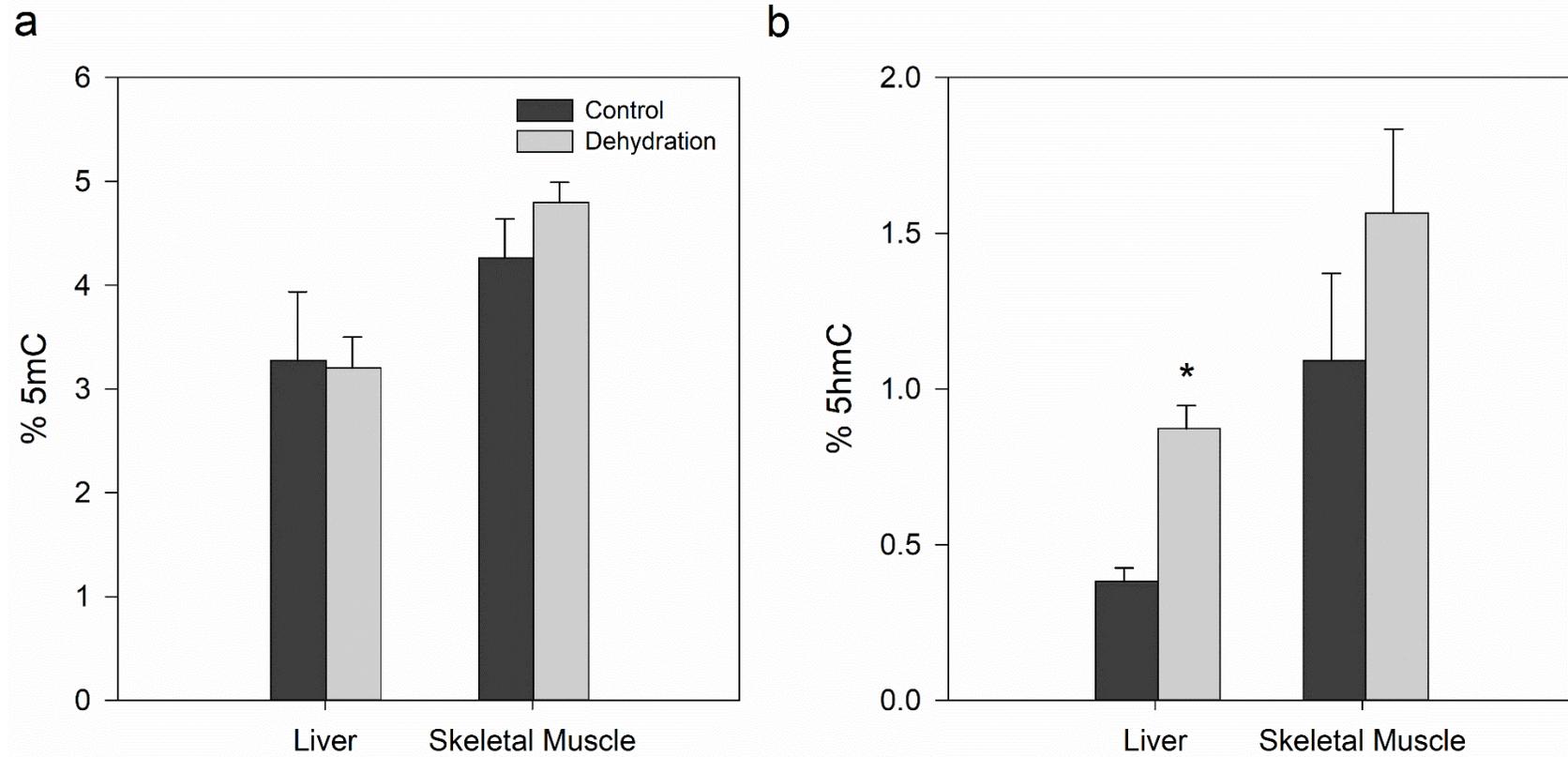


Figure 3.1 Global DNA modification levels in response to dehydration in *Xenopus laevis*. Global levels of a) 5-methylcytosine (5mC) and b) 5-hydroxymethylcytosine (5hmC) were measured in liver and skeletal muscle tissues. Values are mean \pm SEM of modification as percentage of genomic DNA, asterisks indicate statistical significance ($p < 0.05$) using a Student's t-test.

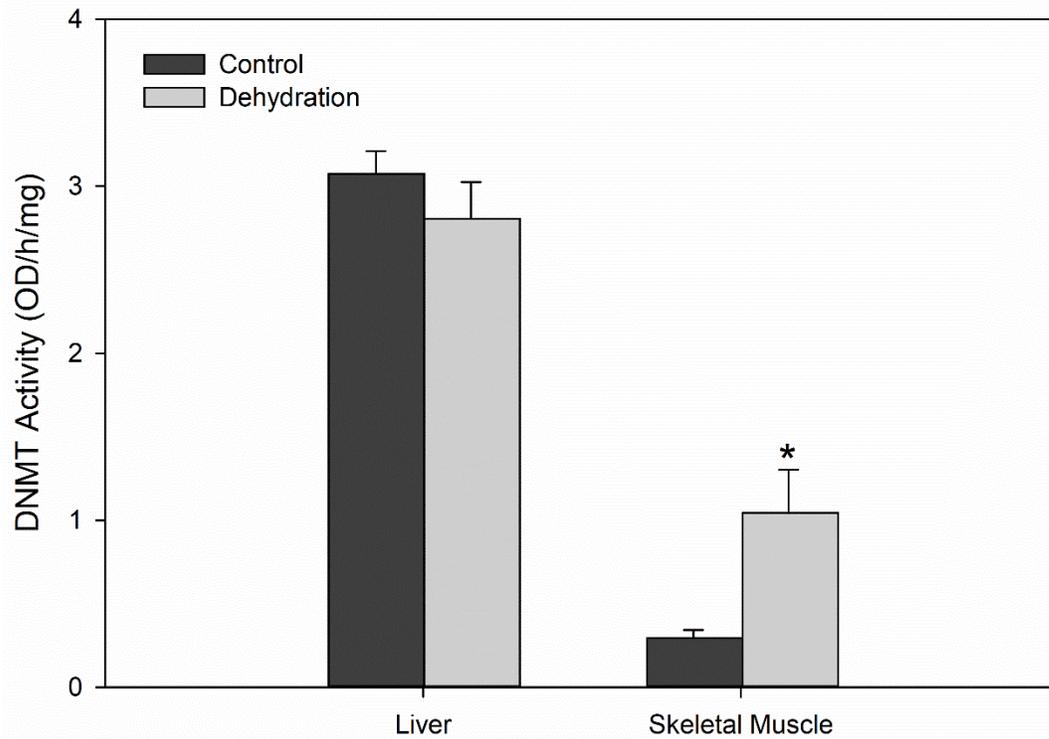


Figure 3.2 DNA methyltransferase activity in response to dehydration in *Xenopus laevis*. Values are mean \pm SEM, asterisks indicate statistical significance ($p < 0.05$) using a Student's t-test.

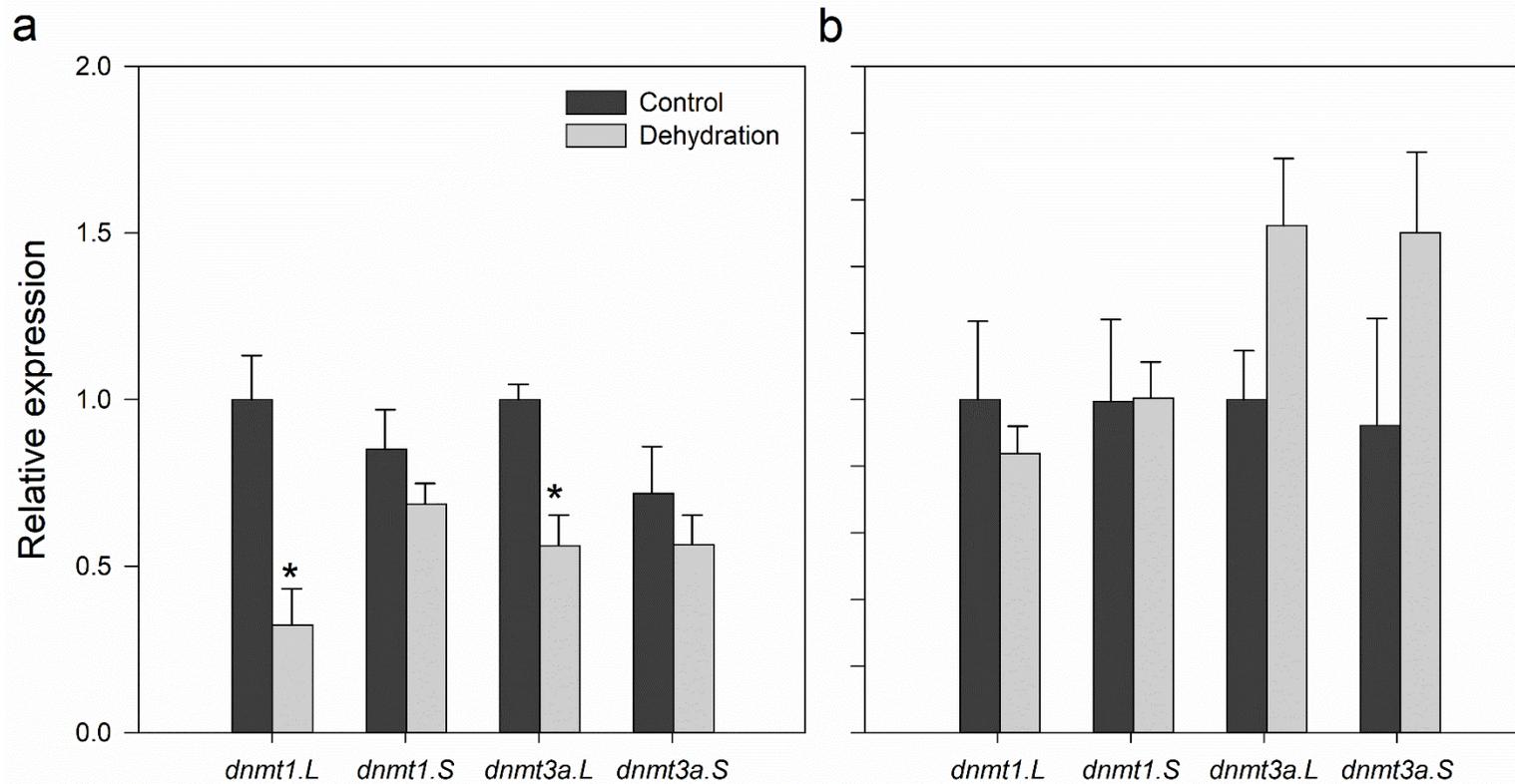


Figure 3.3 Relative expression of *dnmt* genes in *Xenopus laevis* during dehydration. *Dnmt* genes were measured by RT-qPCR in a) liver and b) skeletal muscle. Short chromosome (S) sub-genome gene of each target are relative to the control levels of their respective long chromosome (L) sub-genome gene. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.

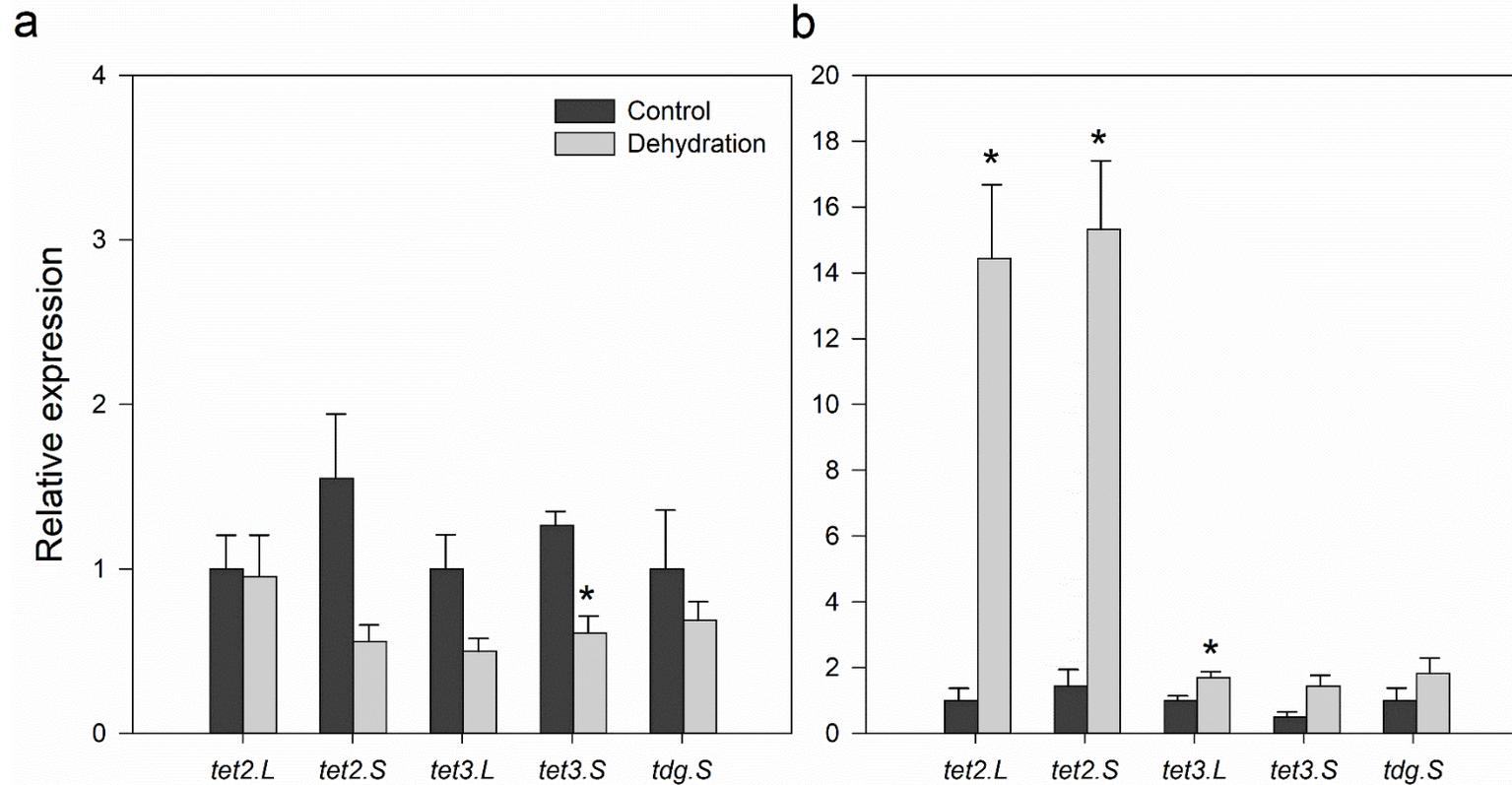


Figure 3.4 Relative expression of DNA demethylation genes in *Xenopus laevis* during dehydration. DNA demethylation genes were measured by RT-qPCR in a) liver and b) skeletal muscle. Short chromosome (S) sub-genome gene of each target are relative to the control levels of their respective long chromosome (L) sub-genome gene. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.

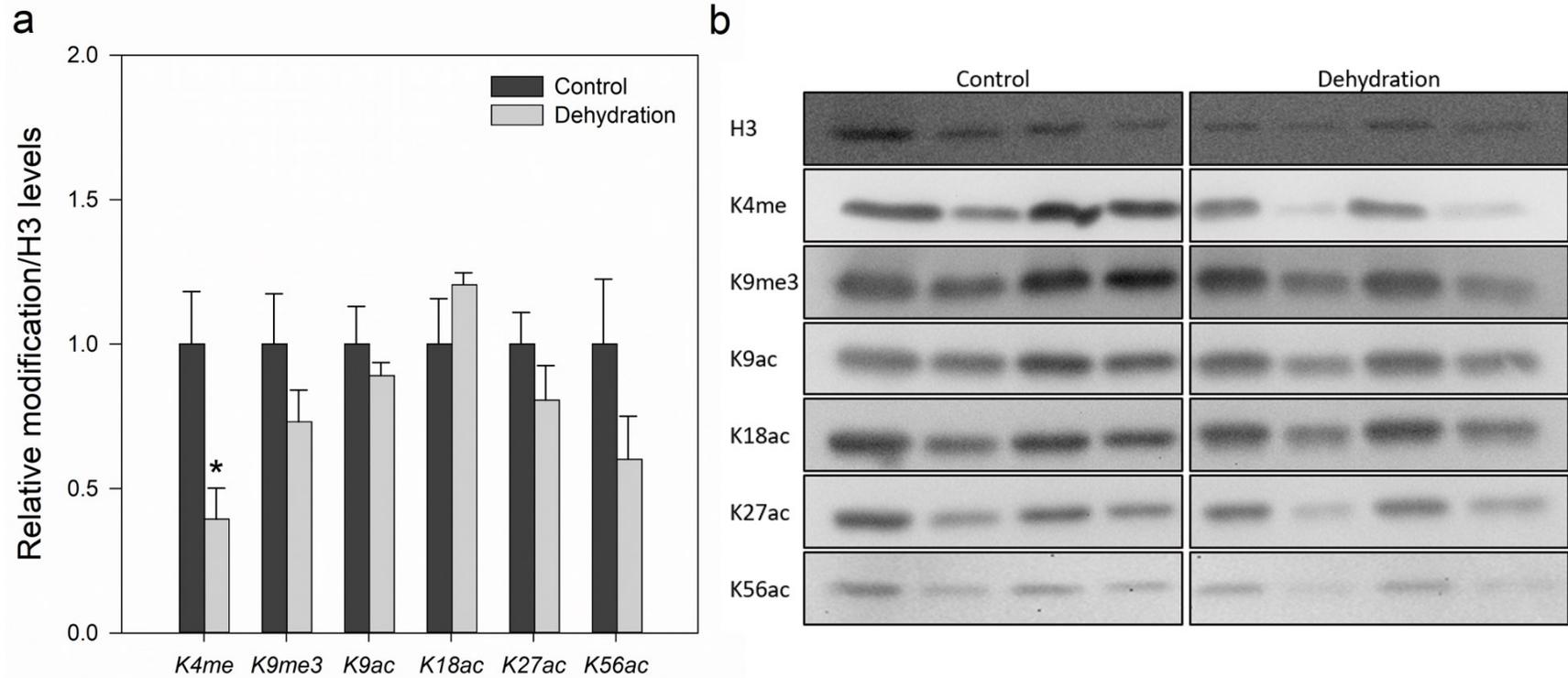


Figure 3.5 Relative changes in histone modifications in the liver of *Xenopus laevis* in response to dehydration. Global levels of histone modifications were measured in the liver using western blotting. All modification levels are relative to total levels of histone H3. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.

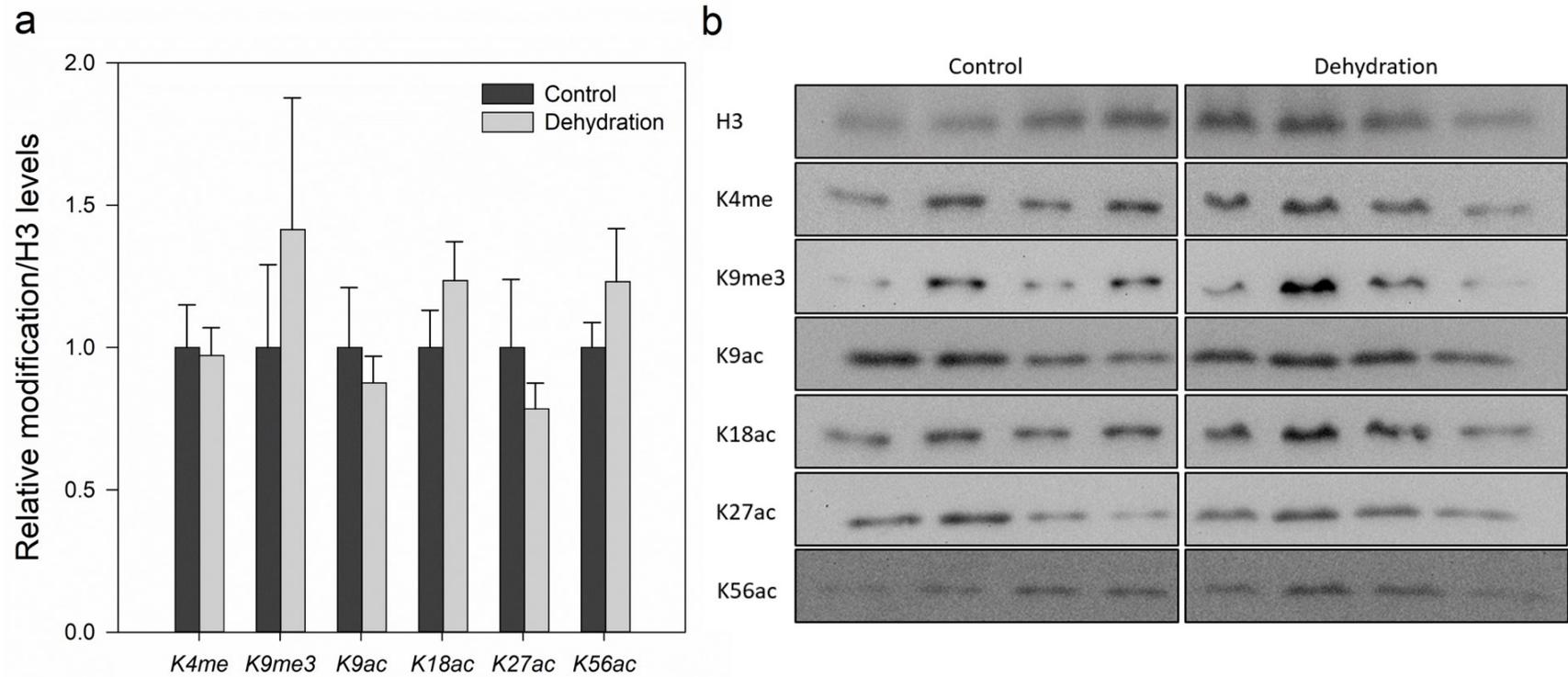


Figure 3.6 Relative changes in histone modifications in the skeletal muscle of *Xenopus laevis* in response to dehydration. Global levels of histone modifications were measured in skeletal muscle using western blotting. All modification levels are relative to total levels of histone H3. Values are mean \pm SEM and asterisks indicate statistical significance ($p < 0.05$) relative to their respective control using a Student's t-test.

Chapter 4

MicroRNA expression in the heart of *Xenopus laevis* facilitates metabolic adaptation to dehydration

Hawkins, Liam J.¹, and Storey, Kenneth B.¹

¹ Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

This study has been published in Genomics:

Hawkins, L. J. & Storey, K. B. (2020). MicroRNA expression in the heart of *Xenopus laevis* facilitates metabolic adaptation to dehydration. Genomics.

<https://doi.org/10.1016/j.ygeno.2020.04.003>

4.1 Abstract

Xenopus laevis survive severe dehydration during the summer months in their natural range. MicroRNA regulate translation of target mRNAs and have shown to be differentially expressed in response to dehydration in *X. laevis*. During dehydration, heart rate is elevated which appears to compensate for the reduced oxygen delivery capability due to increased hematocrit. We hypothesized that microRNAs would be differentially expressed in the heart to modulate gene expression levels in response to dehydration. The present study assessed changes in the microRNAome of *X. laevis* heart in response to severe dehydration (30% loss of body water) using microRNA-seq. We show that target genes are enriched for RNA, DNA, and transcription factor binding activities, cardiac muscle contraction, and glycolysis/gluconeogenesis. These results suggest that microRNAs contribute to gene expression reorganization in the heart in response to dehydration, putatively supporting the increased physiological demands and ATP production needs by the heart under these conditions.

4.2 Introduction

The African clawed frog, *Xenopus laevis*, is well known as an established model organism in the fields of cellular and developmental biology. Comparatively less studied is its natural responses to environmental stresses, especially whole-body water loss that is predicated by the seasonal drying of its normal aquatic environment. Native to southern Africa, *X. laevis* can bury themselves in the mud and initiate multiple tissue specific adaptations as the ponds they inhabit evaporate. The skin secretes a mucus to minimize evaporative water loss, nitrogen waste is shunted from ammonia to urea production (Balinsky et al. 1961), and water is preferentially lost from plasma compared to vital organs, especially the brain (Hillman 1978a; Hillman and Sommerfeldt 1981). As a consequence of reduced plasma volume and increased hematocrit, necessary changes to the cardiorespiratory system can be required and as postulated by Hillman (Hillman 1978b), the critical problem to be overcome is a reduction in circulatory oxygen delivery efficiency, eventually leading to anoxia at the limits of dehydration survival. This is supported by observed increases in whole body lactate under high levels of dehydration in *X. laevis* (Hillman 1978b).

Multiple parameters of the circulatory system are adjusted in response to dehydration, including a linear correlation between heart rate and percent body water loss (Hillman 1978b). *X. laevis* hemoglobin has a higher affinity for O₂ relative to other amphibians, which favours cutaneous uptake, and this affinity further increases during dehydration (Jokumsen and Weber 1980). Their hemoglobin has a high pH sensitivity as well which facilitates unloading in tissues as oxygen becomes limited and P_{CO_2} rises (Jokumsen and Weber 1980). Overall, increased work by the cardiorespiratory system and decreased

oxygen transport would create cellular stresses that necessitate changes to metabolism and gene expression patterns quickly in response to changing environmental conditions. MicroRNA have previously been shown to fit the role of fast-acting regulatory molecules that can respond to severe, acute change to the environment in a variety of animal systems (Biggar and Storey 2018).

MicroRNA are a class of short non-coding RNA, typically ~20-22 nucleotides in length, that act as guides to silence mRNA translation. MicroRNA bind to the 3' - untranslated region (3' -UTR) of mRNAs and depending on the degree of complementarity, this leads to either cleavage degradation of the mRNA (in the case of full complementarity) or translational repression and storage of mRNA transcripts into cellular compartments such as P-bodies (He and Hannon 2004; Liu et al. 2005). MicroRNA are initially transcribed in an immature form known as a primary microRNA (pri-microRNA) (Lee et al. 2002). These structures contain a stem-loop of dsRNA containing what will be the mature sequence, as well as 5' and 3' ssRNA tails. The nuclear ribonuclease Droscha, and its cofactor DGCR8, cleave the pri-microRNA approximately 20-22 bp away from the terminal loop of the hairpin to form a structure known as a precursor microRNA (pre-microRNA) (Han et al. 2004). The pre-microRNA is then translocated out of the nucleus by exportin 5 and RanGTP (Bohnsack et al. 2004). In the cytoplasm a second ribonuclease, Dicer, cleaves the terminal loop from the hairpin to create a 20-22 bp dsRNA molecule (Ketting et al. 2001). This pre-microRNA duplex contains a “guide” strand, which is fully- or partially complimentary to the target mRNA, and a “passenger” strand which is generally not biologically active. One of four members of the Argonaut (AGO) protein family, with the help of the HSC70-HSP90 chaperones,

loads the pre-microRNA duplex, discards the passenger strand and forms the mature RNA-induced silencing complex (RISC) (Iwasaki et al. 2010). Which strand is the guide strand is determined by thermodynamic stability (selecting the strand with the less stable 5' end) and a terminal uracil preference by the AGO protein (Wu et al. 2009; Hu et al. 2009).

The action of microRNA is a powerful post-transcriptional regulatory mechanism. They are metabolically “cheap” to produce, fast-acting, and capable of transcriptome wide targeting. For these reasons it comes as no surprise that microRNAs have been implicated in virtually every cellular and metabolic process including adaptation to adverse environmental conditions. Hence, microRNA regulation is a suitable candidate for gaining information about the molecular mechanisms underlying these adaptations. MicroRNAs have been shown to be differentially regulated in *X. laevis* in response to the loss of 30% of total body water. An initial study of 10 microRNAs in the liver, kidney, and skin tissues showed that three microRNAs decreased during dehydration, and these targeted solute carriers and MAPK signalling pathways (Wu et al. 2013). This was followed up by a study of 43 highly conserved microRNAs in *X. laevis* brain that showed that 12 were decreased during dehydration (Luu and Storey 2015). Pathway analysis of these microRNAs predicted that axon guidance and long-term potentiation pathways were targeted, which aligns with work suggesting that neuroprotective mechanisms are initiated by preferentially maintaining blood flow in the brain (Hillman and Sommerfeldt 1981).

Here we use next-generation sequencing to examine microRNA expression in the heart of *X. laevis* in response to dehydration. We use multiple bioinformatic tools to

assess differential expression of microRNA and gene set enrichment analysis to predict cellular functions and pathways that are under regulation by microRNA. Our results add to our knowledge of adaptations that underpin dehydration tolerance adaptations in the heart.

4.3 Materials and Methods

4.3.1 Animal experiments

Adult male *Xenopus laevis* were procured from the University of Alberta Science Animal Support Services and were acclimated at Carleton University in aerated tanks filled with dechloraminated water for one week at 22°C. Animals were fed three times a week and the water was changed the day of each feeding. The frogs were randomly split into two groups (mean body mass 64.40 ± 10.77 g and 64.58 ± 10.78 g); control and dehydrated. Animals in the control group were kept in the above conditions. Animals in the dehydrated group were weighed and placed in dry containers at 22°C and allowed to lose body water over time. Dehydration group animals were then weighed periodically over the course of two days until they had lost approximately 30% body water to evaporative dehydration (actual % body water loss 33.59 ± 2.58 , dehydration time 25.05 ± 1.98 h). Percent body water loss was calculated using the following equation:

$$\% \text{ Body water loss} = \frac{m_i - m_d}{m_i \times BWC_i} \times 100\%$$

where m_i is the animal initial mass, m_d is the dehydrated mass, and BWC_i is the previously determined initial body water content (Malik and Storey 2009a). Control and dehydrated animals were euthanized by pithing and heart was quickly excised, flash frozen in liquid nitrogen, and stored at -80°C. All protocols were approved by the

Carleton University Animal Care Committee (protocol #106936) and conformed within the guidelines of the Canadian Council on Animal Care.

4.3.2 RNA extraction

RNA was extracted from cardiac tissue of control ($n=3$) and dehydrated ($n=4$) animals. For each sample, 50 mg of frozen tissue was rapidly homogenized using a Polytron homogenizer in 1 mL of TRIzol reagent (Invitrogen; Cat# 15596-018). A 200 μ L aliquot of chloroform was then added to each sample before centrifugation ($10,000 \times g$ at 4°C for 15 min). The upper aqueous phase containing RNA was transferred to a new microtube. The RNA was then precipitated by the addition of 500 μ L of 2-propanol and incubated for 10 min at room temperature. RNA was then pelleted by centrifugation ($10,000 \times g$ at 4°C for 15 min), washed twice with 70% ethanol and air-dried for 10-15 min. RNA pellets were then resuspended in 50 μ L of RNase-free water. RNA concentration and purity were determined using a BioTek Take3 microspot plate and PowerWave HT microplate spectrophotometer. All samples had an OD 260/280 ratio of ~ 2.0 . RNA integrity was determined by electrophoresis on a 1% agarose gel.

4.3.3 Small RNA sequencing

RNA samples were sequenced by the Center for Applied Genomics at the Hospital for Sick Children (Toronto, ON, Canada). RNA quality was validated prior to microRNA library construction using an Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Small RNA cDNA libraries were constructed, validated with the Bioanalyzer, and sequenced using an Illumina HiSeq 2500 platform.

4.3.4 Read processing

Sequencing data for all samples is available in the SRA database (SRA Accession: PRJNA549442). Raw read data was processed as previously described (Zhang et al. 2016). Adapters were trimmed using cutadapt (Martin 2011) and FastQC was used to validate successful adapter removal and distribution of small RNA length (Andrews 2010). To remove reads corresponding to non-microRNA small RNAs (tRNA, rRNA, piRNA, snRNA, and snoRNA) a negative reference file was created from the Rfam (Kalvari et al. 2018) and piRNABank databases (Sai Lakshmi and Agrawal 2008). Bowtie (Langmead et al. 2009) was used to remove reads that aligned to this negative reference index. All reads that did not align were then aligned using bowtie to all mature microRNA sequences from the miRbase microRNA database (Kozomara et al. 2019) with the seed sequence length parameter set to 20 nucleotides and only perfect matches to the seed sequences are reported. The reads that aligned to mature microRNA sequences were then sorted and read counts were determined for each microRNA using samtools (Li et al. 2009) and Unix command line tools. MicroRNAs with less than four reads were removed and read counts were normalized using the voom method (Law et al. 2014) as previously described (Zhang et al. 2016).

4.3.5 Differential expression analysis and clustering

Differential expression of microRNAs between control and dehydrated animals was determined using the limma R package employing linear model fitting with empirical Bayesian testing (Ritchie et al. 2015). Differentially expressed microRNAs were considered significant when the FDR-corrected p-value was less than 0.05 and had a fold-change that satisfied the following equation:

$$|\log_2 FC| \geq \log_2 1.5$$

Significantly differentially expressed microRNAs were clustered hierarchically by microRNA and sample using the Ward method (Ward 1963).

4.3.6 Gene set analysis

Gene set analysis was conducted using the RBiomirGS R package (Zhang and Storey 2018) which implements a logistic regression-based analysis to determine significantly enriched gene ontology (GO) terms (The Gene Ontology Consortium 2019) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa et al. 2019). This package uses multiple databases to enumerate microRNA:mRNA interactions and due to limited *Xenopus laevis* annotations, fully conserved human microRNA orthologs were used. Once microRNA:mRNA interactions are gathered, RBiomirGS uses microRNA expression to determine a microRNA score ($S_{microRNA}$) for each microRNA as previously described (Garcia-Garcia et al. 2016) using the following equation:

$$S_{microRNA} = -\log_{10} p \times \text{signum}(\log_2 FC)$$

where p is the FDR-adjusted p -value and FC is the fold-change of the microRNA between control and dehydrated animals from the differential expression analysis. For each mRNA target a mRNA score (S_{mRNA}) is also calculated as the sign reversed summation of $S_{microRNA}$ of targeting microRNAs using the following equation:

$$S_{mRNA} = -\sum_{i=1}^n S_{microRNA}^i$$

where n is the number of microRNAs targeting the mRNA and $S_{microRNA}^i$ is the $S_{microRNA}$ of the i^{th} targeting microRNA. RBiomirGS then uses the calculated S_{mRNA} of mRNAs to determine significantly enriched gene sets (FDR-adjusted p -value < 0.05) and

corresponding model coefficients for each GO term and KEGG pathway, where a positive model coefficient indicates negative regulation by microRNA is decreased and a negative coefficient indicates negative regulation is increased by microRNA in the dehydration group relative to the control group. Estimated model coefficients and standard error are calculated by a logistic regression based method, which are then used to determine statistical significance (Zhang and Storey 2018).

To further explore GO term enrichment, all GO terms and their corresponding FDR-adjusted p -values were submitted to the REVIGO web server to produce networks of GO terms that are semantically related (Supek et al. 2011). Due to the hierarchical structure of GO terms, GO term enrichment often results in enrichment of numerous redundant terms since both parent and child terms can deliver similar results. This makes biological interpretation more tedious and over-estimates the number of biologically significant results. REVIGO implements an algorithm that accounts for the GO graph structure and reduces functionally redundant terms to summary terms based on semantic similarity (Pesquita et al. 2009; Supek et al. 2011). Enriched GO terms with corresponding FDR-adjusted p -values were used as input to REVIGO with default settings.

4.3.7 Three prime untranslated region analysis

Since RBiomirGS uses human microRNA:mRNA annotations, we validated whether these interactions hold for *X. laevis* using STarMir (Rennie et al. 2014) since this software tool was built from real mRNA:microRNA interaction data and shows excellent cross-species validation (Liu et al. 2013). We gathered 3'UTR sequences for *X. laevis* from Xenbase (Karimi et al. 2018) using the *Xenopus laevis* 9.2 J-Strain genome, and for

human 3'UTR sequences we used the UCSC Genome browser and the hg38 assembly (Lee et al. 2020).

4.3.8 Data visualization

Volcano plots, bar plots, and box plots for differential expression, gene set analysis, and 3'UTR analysis were produced using the matplotlib (Hunter 2007) and seaborn (Waskom et al. 2014) python packages. Visualization of semantically related GO terms produced by REVIGO was done using Cytoscape (Su et al. 2014) and hierarchical clustering of significantly differentially regulated microRNAs was produced by the gplots R package (Warnes et al. 2016).

4.3.9 MicroRNA cDNA synthesis

MicroRNA cDNA synthesis was performed as previously described (Biggar et al. 2014). Briefly, RNA samples were polyadenylated using the Epi-Bio PolyA Tailing Kit (Epicentre, USA; Cat# PAP5104H). Polyadenylation reactions contained 3 µg RNA, 1 µL 10 mM ATP, 1 µL 10x polyadenylate polymerase buffer, 0.5 µL *E. coli* poly (A) polymerase (2U), and RNase-free water to a final volume of 10 µL. Polyadenylation was carried out in an Eppendorf thermocycler by incubating at 37°C for 30 min then terminating the reaction at 95°C for 5 min and immediately chilling on ice. A universal stem loop primer was then annealed to the polyadenylated samples by adding 5 µL of 250 pM universal stem loop primer (5'-CTCACAGTACGTTGGTATCCTTGATGTTTCGATGCCATATTGTACTGTGAGT TTTTTTTTVN-3') to each completed polyadenylation reaction and heating to 95°C for 5 min followed by incubating at 65°C for 5 min and then chilling on ice. The samples were then reverse transcribed using the Invitrogen M-MLV Reverse Transcriptase Kit (Cat#

28025013) with 1 μL of 25 mM dNTPs (BioShop, Burlington, Canada). Samples were then stored at -20°C until use.

4.3.10 RT-qPCR Validation

MicroRNA-seq results were validated using RT-qPCR as previously described (Hawkins et al. 2018). The top 15 significantly differentially regulated microRNAs (by FDR-adjusted p -value) were selected for validation and 11 successfully amplified single products with reaction efficiencies between 90-110%. These microRNAs were used for validation. The five least significantly differentially regulated microRNAs were used as internal controls for normalization as previously described (Hawkins et al. 2018). The use of the most differentially expressed, and least differentially expressed microRNAs in this validation analysis means the results are predicated on differentially expressed microRNA being differentially expressed and non-differentially expressed microRNA being non-differentially expressed. Primer sequences are found in **Supplementary Table 4.1**.

RT-qPCR was performed using a Bio-Rad CFX Connect as previously described (Pellissier et al. 2006). The qPCR reactions contained 8 μL microRNA cDNA, 2 μL qPCR buffer (100 mM Tris-HCl, pH 8.5, 500 mM KCl, 1.5% v:v Triton X-100, and 20 mM MgCl_2), 0.16 μL of 25 mM dNTPs, 4 μL of 1 M trehalose, 0.5 μL of formamide, 0.1 μL of 25x SYBR Green in DMSO (Invitrogen; #S7585), 0.5 μL of each primer (25 nM), 0.125 μL of $5\text{U}\cdot\mu\text{L}^{-1}$ wild-type Taq (BioShop; Cat# TAQ001.1), and autoclaved ddH₂O to a final volume of 20 μL . For each quantification run a 2-fold dilution standard curve was also run to determine reaction efficiency.

The thermocycling parameters were 1) 95°C for 3 min, 2) 40 cycles of 95°C for 10 sec followed by 57°C for 30 sec, and 3) a melt-curve from 65-95°C in 0.5°C for 5 sec increments. The dilution factor of microRNA cDNA for each target was determined by running an initial standard curve for each target. The Pfaffl method (Pfaffl 2001) was used to convert C_q to relative quantities:

$$Q = (E\% \times 0.01 + 1)^{C_{qmin} - C_{qsample}}$$

where Q is the relative quantity, $E\%$ is the efficiency percentage determined from the standard curve, C_{qmin} is the minimum C_q for that target, and $C_{qsample}$ is the C_q of the sample. The relative quantities for the five normalization microRNAs were inputted into the geNorm software package (Vandesompele et al. 2002) to produce normalization factors to normalize the relative quantities of the 11 validation microRNAs.

4.4 Results

4.4.1 Small RNA sequencing summary

Summary results for read processing for each sample are presented in **Supplementary Table 4.2**. Samples from control and dehydration animals averaged $8,516,381 \pm 934,480$ and $8,193,821 \pm 335,749$ raw reads respectively. Trimming and quality filtering resulted in $7,801,107 \pm 864,830$ and $7,563,655 \pm 477,699$ reads remaining for control and dehydration samples. After negative filtering, $1,867,321 \pm 625,947$ and $2,000,352 \pm 182,132$ reads aligned to 202 known microRNA sequences from control and dehydration samples. Adapter ligation bias in small RNA sequencing is known to occur in a sequence-dependent manner (Raabe et al. 2014) and appears to result mostly in under- rather than over-representation of specific microRNA (Fuchs et al. 2015). This means that some low expressed microRNA may not be detected in our

analysis. We also perform analysis using relative expression between control and dehydration samples, rather than comparing expression between microRNA to minimize the effect of this bias.

4.4.2 Differential expression of microRNA in response to dehydration

Of the 202 identified microRNAs in all samples, 24 were significantly differentially regulated between control and dehydrated animals (**Figure 4.1; Supplementary Table 4.3**). Of these differentially regulated microRNAs, 76% were intergenic, 16% were intronic, 4% exonic, and 4% had a genomic position with multiple overlapping genes (**Supplementary Figure 4.1**). This indicates that the expression of the majority of differentially expressed microRNA is likely not a function of a parent gene. Differential expression results indicated that 21 were significantly downregulated during dehydration whereas 3 (miR-25-3p, miR-3529-3p, and miR-7-5p) were significantly upregulated. These results are visualized using a volcano plot where red markers indicate microRNAs that are significantly downregulated and blue markers indicate microRNAs that are significantly upregulated in dehydrated animals (**Figure 4.1**). Differential expression results were validated using RT-qPCR and results from RT-qPCR and small RNA-seq were consistent for all eleven validated microRNAs as described above (**Supplementary Figure 4.2**). Although all validated microRNAs were downregulated in both analyses, the differences in magnitude in some cases may be a result of the ligation bias as previously mentioned, or PCR amplification bias during library preparation for small RNA sequencing that would be accounted by reaction efficiency in the RT-qPCR validation.

Hierarchical clustering analysis of significant differentially expressed microRNA showed all control samples clustered together and all dehydration samples clustered

together (**Figure 4.2**). MicroRNA expression values for each microRNA were standardized to have the same mean and standard deviation giving a z-score. Z-scores are indicated by color where negative z-scores are more orange and positive z-scores are more purple.

4.4.3 Gene ontology terms enriched for differentially expressed microRNA

Gene set enrichment analysis was performed on the GO Molecular Function, Biological Process, and Cellular Component databases. Results showed that 41 GO Molecular Function terms were statistically significantly enriched (FDR-adjusted p -value < 0.05 ; **Supplementary Table 4.4**) with only four having negative model coefficients (*GO Acid Phosphatase Activity*, *GO Racemase and Epimerase Activity*, *GO Phosphatidylinositol Binding*, and *GO Phosphatidylinositol 3-Phosphate Binding*). A positive model coefficient indicates negative regulation by microRNA is decreased and a negative coefficient indicates negative regulation is increased by microRNA in the dehydration group relative to the control group. GO Molecular Function enrichment results are visualized in **Figure 4.3a** and the top 15 most significantly enriched terms (by FDR-adjusted p -value) are shown in **Figure 4.3b**. REVIGO analysis of GO Molecular Function enrichment resulted in four distinct semantically related networks with ≥ 3 members (**Figure 4.4**). Nodes are colored by FDR-adjusted p -value where darker blue corresponds to a smaller p -value, the size of the nodes correspond to the number of semantically related GO terms the node represents, and grey edges between nodes show semantic relationships. The largest network was comprised of two sub-networks (29 and 18 members with one mutual member) corresponding roughly to RNA/DNA/nucleotide binding (**Figure 4.4a**), and protein binding (**Figure 4.4b**). Three of the networks are

small (6-8 members), roughly corresponding to transporter activity (**Figure 4.4c**), post-translational modification and kinase activity (**Figure 4.4d**), and peptidase/hydrolase/phosphatase/nuclease activity (**Figure 4.4e**).

Enrichment analysis of GO Biological Process terms also showed a high skew towards terms having positive model coefficients (**Supplementary Figure 4.3**), where 206 were greater than zero and 9 were less than zero (**Supplementary Table 4.5**). Similar to GO Molecular Function, many of the most significantly enriched GO Biological Process terms involved RNA related processes, for example, *GO mRNA Metabolic Process*, *GO RNA Processing*, *GO nuclear Transcribed mRNA Catabolic Process Nonsense Mediated Decay*, *GO RNA Catabolic Process*, *GO Translational Initiation*, and *GO Ribonucleoprotein Complex Biogenesis* are 5 of the top 10 most enriched processes (**Supplementary Figure 4.3**; **Supplementary Figure 4.6**). REVIGO analysis of GO Biological Process enrichment resulted in a single densely connected network (**Supplementary Figure 4.4**).

GO Cellular Component enrichment showed 35 significantly enriched terms, all of which had positive model coefficients (**Supplementary Figure 4.5**; **Supplementary Table 4.6**). Consistent with the other GO term enrichment results, the most significantly enriched GO Cellular Component terms related to RNA and DNA functions of the cell, in fact 30 of the 35 significantly enriched terms are related to ribosomes, the nucleus, chromosomes, and other DNA and RNA processing components (**Supplementary Table 4.6**). As with GO Biological Process enrichment, REVIGO analysis of GO Cellular Components resulted in a single densely connected network (**Supplementary Figure 4.6**).

4.4.4 KEGG pathways with reduced microRNA regulation

KEGG pathway enrichment was performed in the same manner as GO term enrichment using RBiomirGS. KEGG enrichment analysis showed that 14 KEGG pathways were significantly enriched and all had positive model coefficients (**Figure 4.5; Supplementary Table 4.7**). Particularly interesting, but perhaps unsurprising, is the enrichment of KEGG Cardiac Muscle Contractions (p -value 0.027, model coefficient 0.292). This pathway is visualized in **Figure 4.6** with the microRNAs that target various members of the KEGG pathway, where S_{mRNA} and $S_{microRNA}$ values are indicated by color such that blue is a more positive score and red is a more negative score (full microRNA:mRNA targeting results in **Supplementary Table 4.8**). Due to the relationship between S_{mRNA} and $S_{microRNA}$, if the summation of $S_{microRNA}$ for microRNAs targeting a specific mRNA target is negative (red), the resulting S_{mRNA} would be positive (blue) and vice versa. In total, 57 distinct microRNAs were found to target the 73 members of this KEGG pathway. A tabular version of the KEGG Cardiac Muscle Contraction pathway with all S_{mRNA} and $S_{microRNA}$ is **Supplementary Table 4.9**.

Also enriched is the KEGG Glycolysis Gluconeogenesis pathway (**Figure 4.5b; Supplementary Table 4.10**), which is visualized with targeting microRNAs in **Figure 4.7**. This pathway showed 54 distinct microRNAs targeting the 42 members of the pathway (all S_{mRNA} and $S_{microRNA}$ are available in **Supplementary Table 4.10**). The KEGG targets with the highest summed S_{mRNA} of containing genes (potentially decreased via microRNA silencing during dehydration) include enolase (ENO), pyruvate kinase (PK), hexokinase (HK), lactate dehydrogenase (LDH), and phosphofructokinase 1 (PFK1).

Since the microRNA:mRNA annotations are derived from human data in the RBiomirGS package, we compared the predicted microRNA:mRNA interactions from human microRNA and 3'UTR sequences with those from *X. laevis* for all genes and microRNA in the *Glycolysis Gluconeogenesis* pathway (**Supplementary Figure 4.7**). We found that in all cases, except for triosephosphate isomerase (TPI), the mean logistic probability of interaction sites was equal to or higher for *X. laevis* than human microRNA:mRNA pairs (**Supplementary Figure 4.7a**), and for 12 of 17 genes the maximum logistic probability was higher for *X. laevis* than human microRNA:mRNA pairs, including TPI (**Supplementary Figure 4.7b**). These findings suggest that although 3'UTR sequences are less conserved than microRNA sequences, their interactions are conserved or even strengthened.

4.5 Discussion

Protein levels depend on multiple factors including mRNA transcription, mRNA sequestering, mRNA stability, and the translational machinery, all regulated by a host of upstream mechanisms. MicroRNA comprise just one component of translational regulation, but the efficiency and speed with which this type of translational regulation can affect protein levels make microRNA particularly suited for adaptive responses to stresses or harsh environmental conditions (Kucherenko and Shcherbata 2018; Biggar and Storey 2018).

4.5.1 Most differentially expressed microRNA downregulated during dehydration

MicroRNA in the heart of *X. laevis* responded to dehydration with a general trend of decreased expression (**Figure 4.1**). Since microRNA are negative regulators of translation this would suggest there is a net reduction of microRNA-mediated

translational inhibition in response to dehydration. It has been suggested that during dehydration exposure, an increased hematocrit will limit oxygen delivery to tissues which can be overcome by increasing the work done by the circulatory system (Hillman 1980). This is apparent in *X. laevis* by a demonstrated linear increase of heart rate with increasing evaporative dehydration (Hillman 1978b). These conditions produce a stressful cellular environment as evidenced by increases in antioxidant expression (Malik and Storey 2009b) and extracellular signalling cascades (Malik and Storey 2009a) in the dehydrating animal. The decreased levels of many microRNAs seen in the heart may therefore contribute to the activation of various cellular responses to dehydration that are necessary for survival. Similar trends were seen in the brain of *X. laevis* during dehydration (Luu and Storey 2015). Analysis of the downregulated microRNAs in the brain in response to dehydration suggested that neuroprotective pathways were being activated. The heart and brain are similar in this regard since their integrity and function during bouts of dehydration exposure is crucial both during the stress and afterwards. Therefore, it may be the case that the downregulation of microRNA in the heart is effectively increasing the translation of cardio- and cytoprotective proteins.

The microRNA showing the most significant differential expression during dehydration was miR-99b-5p (**Figure 4.1**), which was expressed at just 15% of control values. While the exact function of miR-99b-5p during dehydration is unknown, this microRNA shows the opposite expression pattern in multiple human heart diseases including dilated cardiomyopathy and aortic stenosis (Ikeda et al. 2007). Both diseases involve a reduction in the ability of the heart to pump blood and, hence, there is a

possibility that a reduction in miR-99b-5p expression, as seen in *X. laevis*, may be related to enhancing cardiac output.

When we performed hierarchical clustering on the significantly differentially expressed microRNAs, the animals from each experimental condition clustered together at the highest level of the dendrogram (**Figure 4.2**). This is not surprising given that most microRNAs showed decreased expression during dehydration, but this adds further evidence that the patterns of microRNA expression at some level are a coordinated response to dehydration and not just stochastic differences in expression between groups. Furthermore, these microRNAs may also be used to differentiate control from dehydration exposed animals, meaning that some might potentially be used as biomarkers for dehydration, although the impracticality of using heart tissue for this purpose is obvious.

4.5.2 RNA and DNA binding targets highly enriched during dehydration

After examining the general expression patterns and highly differentially expressed microRNAs, we performed gene set enrichment analysis on GO gene sets and KEGG pathways. The RBiomirGS program (Zhang and Storey 2018) produces a model coefficient for each analyzed gene set, which is a measure of how microRNAs targeting that set of genes changes between conditions. For example, if all microRNAs targeting a specific pathway increase during dehydration, the microRNAs would contribute negative regulation to the members of that pathway and the model coefficient would be negative to reflect this probable downregulation. The reverse is also true, if all microRNAs targeting a pathway decrease, a positive model coefficient is produced to reflect reduced potential negative regulation by those microRNAs. Consistent with our results showing a skew

towards decreased microRNA expression during dehydration, most significantly enriched GO molecular function terms have a positive model coefficient (**Figure 4.3**). Although there are four GO molecular function terms with negative model coefficients, they appear semantically unrelated to the terms with positive coefficients. The terms with positive coefficients are commonly related to RNA, DNA, and transcription factor binding, whereas the terms with negative coefficients involve phosphatase and racemase activities and phosphatidylinositol binding (**Supplementary Table 4.4**). The semantic partitioning between positive and negative coefficient terms further validates the biological relevance of this type of analysis. It is expected that when encountering an environmental stress such as dehydration, *X. laevis* would activate specific cellular processes while turning off others, and this is reflected in the microRNA data presented here.

The potential removal of microRNA mediated negative regulation is reflected in the semantic relation networks produced using REVIGO (**Figure 4.4**). The largest semantically related network of enriched GO molecular function terms during dehydration contains two sub-networks that relate to DNA/RNA binding (**Figure 4.4a**) and protein/transcription factor binding (**Figure 4.4b**), respectively. This suggests that transcription and translation are specifically regulated by changes in microRNA levels, which would contribute to gene expression reprogramming during dehydration. These results are consistent with previous studies showing that second messenger signalling cascades and heat shock proteins have altered expression in response to dehydration in this species (Malik and Storey 2009a; Luu et al. 2018). The most significantly enriched nodes were poly(A) RNA-, RNA-, and mRNA-binding, which suggests microRNAs are affecting proteins that are also involved in post-transcriptional regulation. RNA binding

proteins influence a variety of mRNA parameters including stability, subcellular localization, transcription, and translation, all of which influence protein levels (Glisovic et al. 2008). The necessity of reorganizing gene expression to survive dehydration is obvious and the efficiency of microRNA make them candidates to contribute to this process. MicroRNA can affect translation quickly enough to react to increasing dehydration by facilitating increases in cytoprotective mechanisms, metabolic reorganization, and more. Three small networks were also identified as being targeted by microRNA: transporter and channel activity (**Figure 4.4c**), post-translational modification related activity (**Figure 4.4d**), and activities related to cleavage of molecular bonds (phosphatase, helicase, nuclease, peptidase, hydrolase etc.)(**Figure 4.4e**). Dehydration necessarily changes solute concentrations and, therefore, distributions of solutes (e.g. ions, metabolites, osmolytes, etc.) become perturbed if no active measures are taken by cells and tissues. Regulating transporter and channel proteins by changing microRNAs could therefore help to adjust solute distributions when this animal is faced with dehydration. It is particularly important in the heart to achieve this given its reliance on ion concentrations to perform muscle contraction. The small networks that show post-translational modification and cleavage activities also broadly tie into the largest network of DNA, RNA, and transcription factor binding. While members of the large network affect which genes are transcribed and translated, post-translational modification and cleavage control the functions of proteins and macromolecules that are already present in the cell. Together these networks show that of all possible molecular functions in the cell, those that affect gene expression and activity are the primary targets of microRNAs in response to dehydration.

Similar to enriched GO terms, all significantly enriched KEGG pathways have positive model coefficients which is due to the skew towards decreased microRNA expression during dehydration (**Figure 4.5**). Consistent with GO term enrichment results indicating decreased microRNA regulation of translation related genes and functions, the *KEGG Ribosome pathway* had the highest model coefficient. This adds further evidence that the response to dehydration involves post-transcriptional regulation of gene expression. Also important to the physiology of cardiac tissue are cell-cell junctions. As seen in our results, the *KEGG Adherens Junction* and *KEGG Focal Adhesion* pathways have positive model coefficients. The elevated heart rate exhibited during dehydration in *X. laevis* may be necessitating regulation of these pathways to affect the electromechanical properties between adjacent cardiomyocytes (Hillman 1978b; Takada et al. 2007). MicroRNA mediated regulation of these pathways may also serve to increase hemodynamic sensitivity which is crucial to implementing adaptive responses in the heart (Cooper 1997; Domingos et al. 2002), and may be important in responding to dehydration.

4.5.3 MicroRNAs targeting cardiac muscle contraction are downregulated

Relevant to what we know about the physiology of dehydration tolerance in this animal, it is not surprising that the *Cardiac Muscle Contraction* KEGG pathway is one of the most significantly enriched pathways (**Figure 4.6**). Since heart rate is elevated during dehydration, the removal of negative regulation of many genes in this pathway may support this increased cardiac function for the duration of dehydration exposure. Myosin and tropomyosin (TPM) were the only myofibril genes targeted by the microRNAs identified here. Myosin binding to the actin filament is highly regulated – tropomyosin

being an integral component of this process (Ehler 2016). Changes in the amount of these proteins may adjust the parameters of the cross-bridge cycle to sustain the needed cardiac output during dehydration, however future experiments will have to explore this point.

The majority of ion transporter proteins involved in the *Cardiac Muscle Contraction* KEGG pathway showed changes in targeting microRNAs associated with these proteins, including SERCA2a and RyR2 (**Figure 4.6**). SERCA2a is the calcium ATPase responsible for pumping calcium into the sarcoplasmic reticulum, whereas the RyR2 channel allows for its release, both of which function to complete cardiomyocyte calcium cycling that enables cardiac muscle contraction. In many cases irregular calcium cycling results in heart failure, and reduced expression of SERCA2a both at the mRNA and protein level is associated with reduced cardiac function (Park and Oh 2013). In fact, supplemental expression of SERCA2a delivered by adeno-associated virus can rescue cardiac function both in patients with end-stage heart failure and in animal models of heart failure (Kawase et al. 2008; Jessup et al. 2011). This could suggest that the reduction in microRNAs targeting SERCA2a as seen in our results could lead to increased SERCA2a levels and heightened cardiac function during dehydration (**Figure 4.6**).

4.5.4 Downregulation of microRNAs targeting most glycolytic enzymes

The *Glycolysis Gluconeogenesis* KEGG pathway was also significantly enriched during dehydration in the heart. This is particularly relevant to the body of literature examining glycolysis and glucose metabolism in *X. laevis* and its contribution to dehydration tolerance. Early studies indicated that at high levels of evaporative dehydration, whole body lactate increased indicating enhanced anaerobic glycolysis in

tissues(Hillman 1978b). This is supported by increases in plasma glucose, presumably from the liver, that can be taken up by tissues such as the heart to fuel anaerobic ATP production (Malik and Storey 2009a). In *X. laevis* skeletal muscle, hexokinase showed decreased affinity for glucose, which was postulated to regulate the entry of glucose into glycolysis when plasma glucose levels are high and skeletal muscle function is not prioritized (Childers and Storey 2016). A few studies have also examined glycolytic enzymes in the liver. Pyruvate kinase transcripts increased during dehydration, and affinity for phosphoenolpyruvate increased which would facilitate this rate-limiting, ATP-producing, step of glycolysis (Dawson et al. 2018). Similarly, LDH from dehydrated animals showed modified kinetic properties in the face of the increased urea produced during dehydration, which was not the case for LDH purified from control animals (Katzenback et al. 2014). In the heart, little is known about glycolysis within the context of *X. laevis* dehydration tolerance. Some glycolytic protein levels have been measured in the heart (Wu et al. 2017), however, the limitations of available reagents meant that glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a normalizing factor and may confound measurement of other glycolytic enzymes if protein levels across the pathway trend up or down. Presumably glycolysis increases in the heart to meet increased ATP requirements to sustain an increased heart rate during dehydration when oxygen is limited. Our results support this hypothesis and show that most enzymes in the Glycolysis Gluconeogenesis KEGG pathway have a positive S_{mRNA} . This means that the microRNAs targeting these enzymes would decrease during dehydration, lifting negative translational regulation of these enzymes. Interestingly, the rate-limiting regulatory enzymes of glycolysis (HK, PFK1, and PK) were among the

enzymes with the highest S_{mRNA} (**Figure 4.7; Supplementary Table 4.10**). Regulating these enzymes post-transcriptionally via microRNA would be the most efficient way for microRNAs to contribute to increasing glycolytic flux. Also predicted to have a reduction in negative regulation is LDH, which when combined with increasing microRNAs targeting of pyruvate dehydrogenase E3 and phosphoglucomutase to potentially block entry to the TCA cycle and glycogen metabolism respectively, paints a picture of glucose flux towards lactate in so far as it can be dictated by translational regulation. While metabolic flux analysis and study of glycolytic enzyme kinetics will be needed to make conclusions about this pathway, our results show that glycolysis in the heart may be supported by post-transcriptional regulation by microRNA in response to dehydration.

4.5.5 Conclusions

The results presented here show that the microRNAome in the heart of *X. laevis* responds to dehydration by generally decreasing microRNA expression. This downregulation would relieve negative regulation of many pathways and cellular process and contribute to the reorganization of gene expression necessary for dehydration tolerance. Gene set enrichment analysis indicates that myriad proteins with RNA, DNA, and transcription factor binding activities have reduced microRNA regulation which may reorganize gene expression programs to quickly respond to increasing dehydration levels. Cardiac muscle contraction itself was also significantly enriched for regulation by microRNA, most of which involved microRNAs that target mRNAs encoding transmembrane transporter proteins. This is perhaps unsurprising given the central need to control ion concentrations to retain muscle contractility and the connection between dehydration and increasing solute concentrations, especially salt ions. The data also

reveal a microRNA mediated post-transcriptional contribution to increasing anaerobic glycolysis. This would help facilitate energy production needed to sustain an increased heart rate even as dehydration begins to increase blood viscosity and impair oxygen delivery. Overall, changes in microRNA expression appear to contribute to gene expression, physiology, and biochemistry of the heart when encountering dehydration.

4.6 References

- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Biggar KK, Storey KB (2018) Functional impact of microRNA regulation in models of extreme stress adaptation. *J Mol Cell Biol* 10:93–101
- Biggar KK, Wu C-W, Storey KB (2014) High-throughput amplification of mature microRNAs in uncharacterized animal models using polyadenylated RNA and stem-loop reverse transcription polymerase chain reaction. *Anal Biochem* 462:32–4
- Bohnsack MT, Czaplinski K, Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 10:185–191 . doi: 10.1261/rna.5167604
- Childers CL, Storey KB (2016) Post-translational regulation of hexokinase function and protein stability in the aestivating frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Cooper G (1997) Basic determinants of myocardial hypertrophy: a review of molecular mechanisms. *Annu Rev Med* 48:13–23 . doi: 10.1146/annurev.med.48.1.13
- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- Domingos PP, Fonseca PM, Nadruz W, Franchini KG (2002) Load-induced focal adhesion kinase activation in the myocardium: role of stretch and contractile activity. *Am J Physiol Heart Circ Physiol* 282:H556–64 . doi: 10.1152/ajpheart.00534.2001
- Ehler E (2016) Cardiac cytoarchitecture — why the “hardware” is important for heart function! *Biochim Biophys Acta - Mol Cell Res* 1863:1857–1863 . doi: 10.1016/J.BBAMCR.2015.11.006
- Fuchs RT, Sun Z, Zhuang F, Robb GB (2015) Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. *PLoS One* 10:e0126049 . doi: 10.1371/journal.pone.0126049
- Garcia-Garcia F, Panadero J, Dopazo J, Montaner D (2016) Integrated gene set analysis for microRNA studies. *Bioinformatics* 32:2809–2816 . doi: 10.1093/bioinformatics/btw334
- Glisovic T, Bachorik JL, Yong J, Dreyfuss G (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 582:1977–1986 . doi: 10.1016/j.febslet.2008.03.004
- Han J, Lee Y, Yeom K, et al (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18:3016–27 . doi: 10.1101/gad.1262504

- Hawkins LJ (2020) Supplementary tables for Chapter 4 of doctoral thesis by Liam Hawkins, Carleton University, 2020. Harvard Dataverse. V1. UNF:6:EEKgFonlctob81SDbHK0A==. doi: 10.7910/DVN/CBNUUF
- Hawkins LJ, Luu BE, Storey KB (2018) Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. *Gene Reports* 13:192–198 . doi: 10.1016/j.genrep.2018.10.006
- He L, Hannon GJ (2004) MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5:522–31 . doi: 10.1038/nrg1379
- Hillman SS (1978a) Some effects of dehydration on internal distributions of water and solutes in *Xenopus laevis*. *Comp Biochem Physiol Part A Physiol* 61:303–307
- Hillman SS (1978b) The roles of oxygen delivery and electrolyte levels in the dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Hillman SS (1980) Physiological correlates of differential dehydration tolerance in anuran amphibians. *Copeia* 1980:125–129
- Hillman SS, Sommerfeldt RW (1981) Microsphere studies of amphibian systemic blood flow redistribution during dehydration, hypovolemia, and salt load. *J Exp Zool* 218:305–308
- Hu HY, Yan Z, Xu Y, et al (2009) Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* 10:413 . doi: 10.1186/1471-2164-10-413
- Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9:90–95 . doi: 10.1109/MCSE.2007.55
- Ikeda S, Kong SW, Lu J, et al (2007) Altered microRNA expression in human heart disease. *Physiol Genomics* 31:367–73 . doi: 10.1152/physiolgenomics.00144.2007
- Iwasaki S, Kobayashi M, Yoda M, et al (2010) Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Mol Cell* 39:292–299 . doi: 10.1016/j.molcel.2010.05.015
- Jessup M, Greenberg B, Mancini D, et al (2011) Calcium Upregulation by Percutaneous Administration of Gene Therapy in Cardiac Disease (CUPID): a phase 2 trial of intracoronary gene therapy of sarcoplasmic reticulum Ca²⁺-ATPase in patients with advanced heart failure. *Circulation* 124:304–13 . doi: 10.1161/CIRCULATIONAHA.111.022889
- Jokumsen A, Weber RE (1980) Haemoglobin-oxygen binding properties in the blood of *Xenopus laevis*, with special reference to the influences of aestivation and of temperature and salinity acclimation. *J Exp Biol* 86:19–37
- Kalvari I, Argasinska J, Quinones-Olvera N, et al (2018) Rfam 13.0: Shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res* 46:D335–D342 . doi: 10.1093/nar/gkx1038
- Kanehisa M, Sato Y, Furumichi M, et al (2019) New approach for understanding genome variations in KEGG. *Nucleic Acids Res* 47:D590–D595 . doi: 10.1093/nar/gky962

- Karimi K, Fortriede JD, Lotay VS, et al (2018) Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res* 46:D861–D868 . doi: 10.1093/nar/gkx936
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Kawase Y, Ly HQ, Prunier F, et al (2008) Reversal of cardiac dysfunction after long-term expression of SERCA2a by gene transfer in a pre-clinical model of heart failure. *J Am Coll Cardiol* 51:1112–9 . doi: 10.1016/j.jacc.2007.12.014
- Ketting RF, Fischer SEJ, Bernstein E, et al (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15:2654–9 . doi: 10.1101/gad.927801
- Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) MiRBase: From microRNA sequences to function. *Nucleic Acids Res* 47:D155–D162 . doi: 10.1093/nar/gky1141
- Kucherenko MM, Shcherbata HR (2018) miRNA targeting and alternative splicing in the stress response - events hosted by membrane-less compartments. *J Cell Sci* 131:jcs202002
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: . doi: 10.1186/gb-2009-10-3-r25
- Law CW, Chen Y, Shi W, Smyth GK (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15:R29 . doi: 10.1186/gb-2014-15-2-r29
- Lee CM, Barber GP, Casper J, et al (2020) UCSC Genome Browser enters 20th year. *Nucleic Acids Res* 48:D756–D761 . doi: 10.1093/nar/gkz1012
- Lee Y, Jeon K, Lee J-T, et al (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* 21:4663–70 . doi: 10.1093/emboj/cdf476
- Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079 . doi: 10.1093/bioinformatics/btp352
- Liu C, Mallick B, Long D, et al (2013) CLIP-based prediction of mammalian microRNA binding sites. *Nucleic Acids Res* 41: . doi: 10.1093/nar/gkt435
- Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7:719–723 . doi: 10.1038/ncb1274
- Luu BE, Storey KB (2015) Dehydration triggers differential microRNA expression in *Xenopus laevis* brain. *Gene* 573:64–9
- Luu BEE, Wijenayake S, Malik AII, Storey KBB (2018) The regulation of heat shock proteins in response to dehydration in *Xenopus laevis*. *Cell Stress Chaperones* 23:45–53 . doi: 10.1007/s12192-017-0822-9

- Malik AI, Storey KB (2009a) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Malik AI, Storey KB (2009b) Activation of antioxidant defense during dehydration stress in the African clawed frog. *Gene* 442:99–107 . doi: 10.1016/j.gene.2009.04.007
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10–12
- Park WJ, Oh JG (2013) SERCA2a: A prime target for modulation of cardiac contractility during heart failure. *BMB Rep* 46:237–243 . doi: 10.5483/BMBRep.2013.46.5.077
- Pellissier F, Glogowski CM, Heinemann SF, et al (2006) Lab assembly of a low-cost, robust SYBR green buffer system for quantitative real-time polymerase chain reaction. *Anal Biochem* 350:310–2 . doi: 10.1016/j.ab.2005.12.002
- Pesquita C, Faria D, Falcão AO, et al (2009) Semantic Similarity in Biomedical Ontologies. *PLoS Comput Biol* 5:e1000443 . doi: 10.1371/journal.pcbi.1000443
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45 . doi: 10.1093/nar/29.9.e45
- Raabe CA, Tang T-H, Brosius J, Rozhdestvensky TS (2014) Biases in small RNA deep sequencing data. *Nucleic Acids Res* 42:1414–26 . doi: 10.1093/nar/gkt1021
- Rennie W, Liu C, Carmack CS, et al (2014) STarMir: A web server for prediction of microRNA binding sites. *Nucleic Acids Res* 42:114–118 . doi: 10.1093/nar/gku376
- Ritchie ME, Phipson B, Wu D, et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47 . doi: 10.1093/nar/gkv007
- Sai Lakshmi S, Agrawal S (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res* 36:D173-7 . doi: 10.1093/nar/gkm696
- Su G, Morris JH, Demchak B, Bader GD (2014) Biological network exploration with Cytoscape 3. *Curr Protoc Bioinforma* 47:8.13.1-24 . doi: 10.1002/0471250953.bi0813s47
- Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800 . doi: 10.1371/journal.pone.0021800
- Takada Y, Ye X, Simon S (2007) The integrins. *Genome Biol* 8:215 . doi: 10.1186/gb-2007-8-5-215
- The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* 47:D330–D338 . doi: 10.1093/nar/gky1055
- Vandesompele J, De Preter K, Pattyn F, et al (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034
- Ward JHJ (1963) Hierarchical grouping to optimize an objective function. *Am Stat Assoc* 58:236–244

- Warnes GR, Bolker B, Bonebakker L, Gentleman R (2016) Package ‘gplots’. Various R Programming Tools for Plotting Data
- Waskom M, Botvinnik O, Hobson P, et al (2014) Seaborn: statistical data visualization
- Wu C-W, Tessier SN, Storey KB (2017) Regulation of the insulin-Akt signaling pathway and glycolysis during dehydration stress in the African clawed frog *Xenopus laevis*. *Biochem Cell Biol* 95:663–671 . doi: 10.1139/bcb-2017-0117
- Wu CW, Biggar KK, Storey KB (2013) Dehydration mediated microRNA response in the African clawed frog *Xenopus laevis*. *Gene* 529:269–75
- Wu H, Ye C, Ramirez D, Manjunath N (2009) Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS One* 4: . doi: 10.1371/journal.pone.0007566
- Zhang J, Hadj-Moussa H, Storey KB (2016) Current progress of high-throughput microRNA differential expression analysis and random forest gene selection for model and non-model systems: an R implementation. *J Integr Bioinform* 13:306
- Zhang J, Storey KB (2018) RBiomirGS: an all-in-one miRNA gene set analysis solution featuring target mRNA mapping and expression profile integration. *PeerJ* 6:e4262

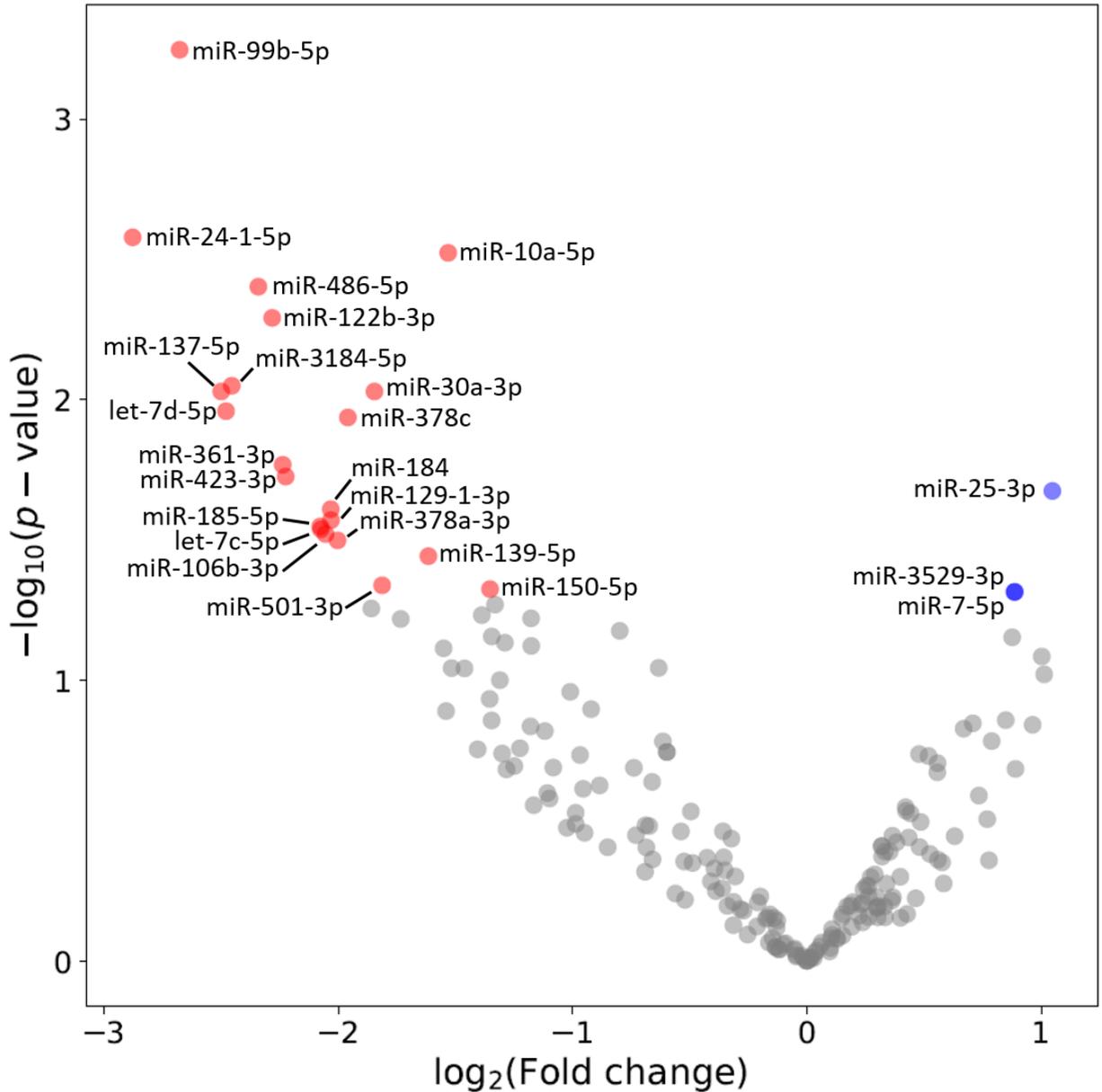


Figure 4.1 MicroRNA differentially expressed in the heart of *Xenopus laevis* in response to dehydration. Volcano plot fold-change thresholds were set to $\pm \log_2 1.5$ and an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate microRNA that were significantly downregulated and upregulated, respectively and grey circles are microRNA that did not pass the fold-change and p -value thresholds

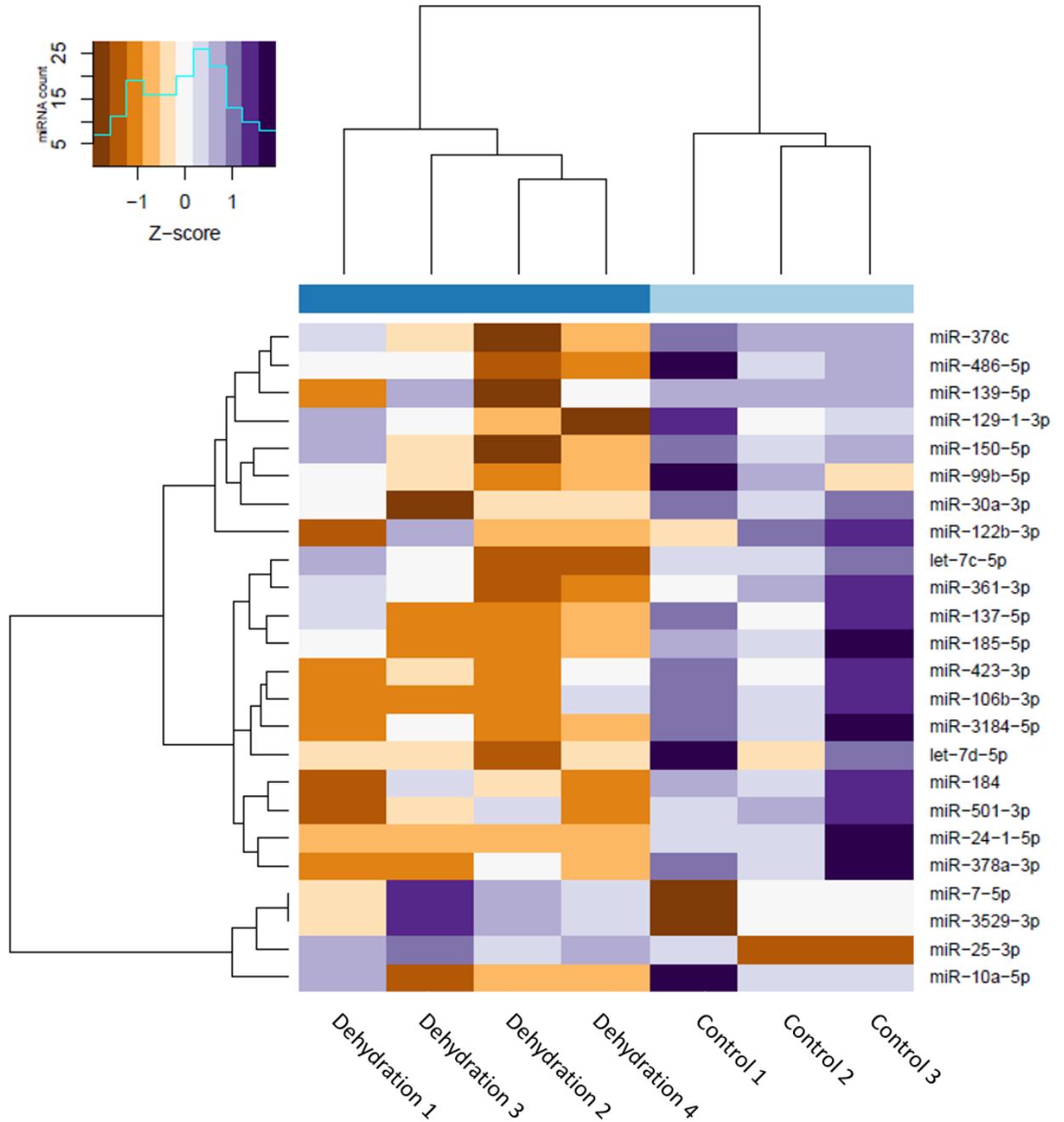


Figure 4.2 Hierarchical clustering of significant differentially expressed microRNA in the heart of *Xenopus laevis* in response to dehydration. Clustering of samples and microRNA was done using the Ward method. Purple and orange indicate higher and lower z-scores (standardized expression level) for each microRNA in each sample. Clustering of samples at the highest level of the dendrogram is indicated by dark and light blue bar above the heat map

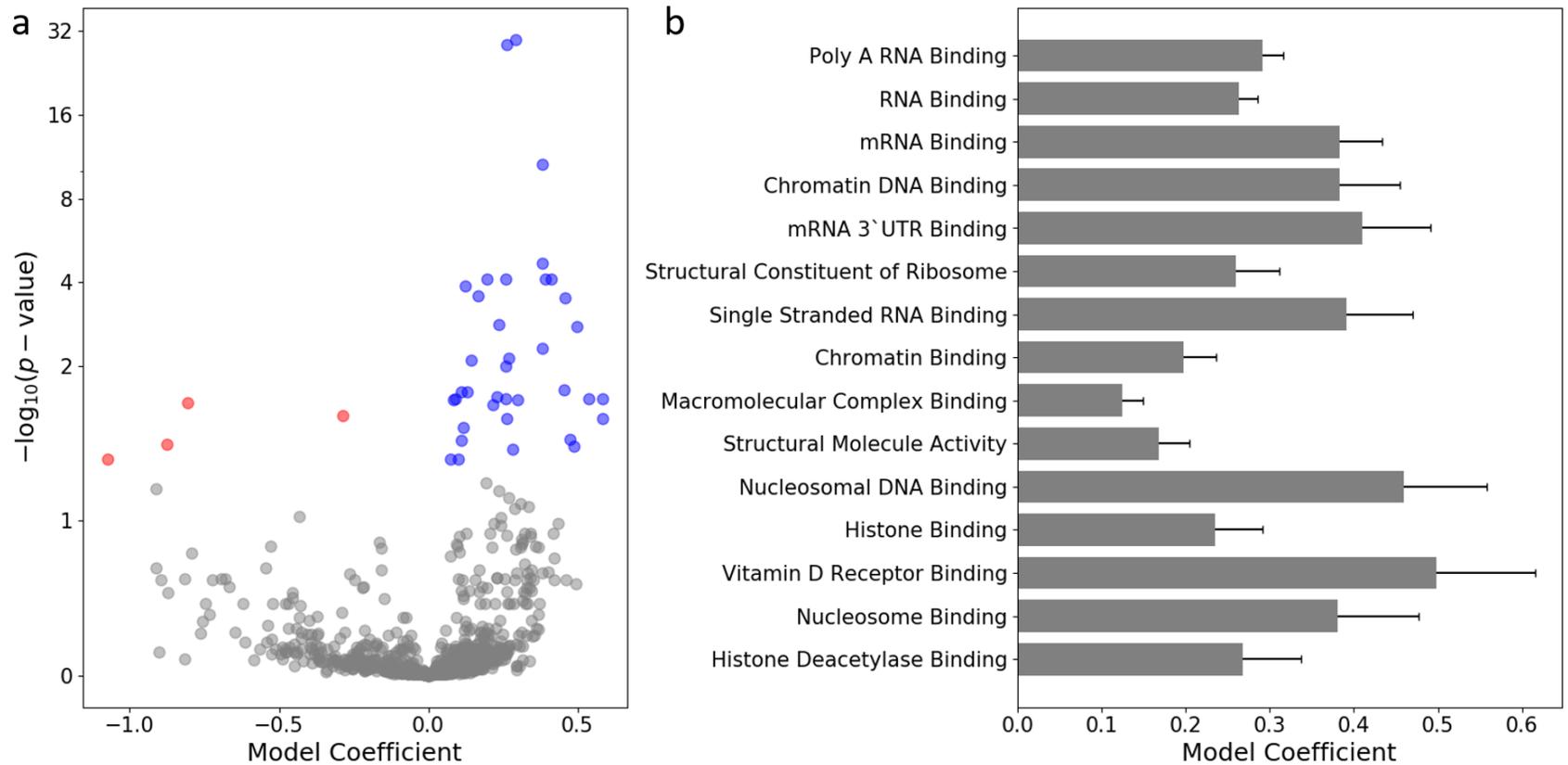


Figure 4.3 Enrichment of GO Molecular Function terms in heart of dehydrated *Xenopus laevis*. **a)** Volcano plot of GO Molecular Function term enrichment with an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. **b)** Top 15 most significantly enriched (by FDR-adjusted p -value) GO Molecular Function terms

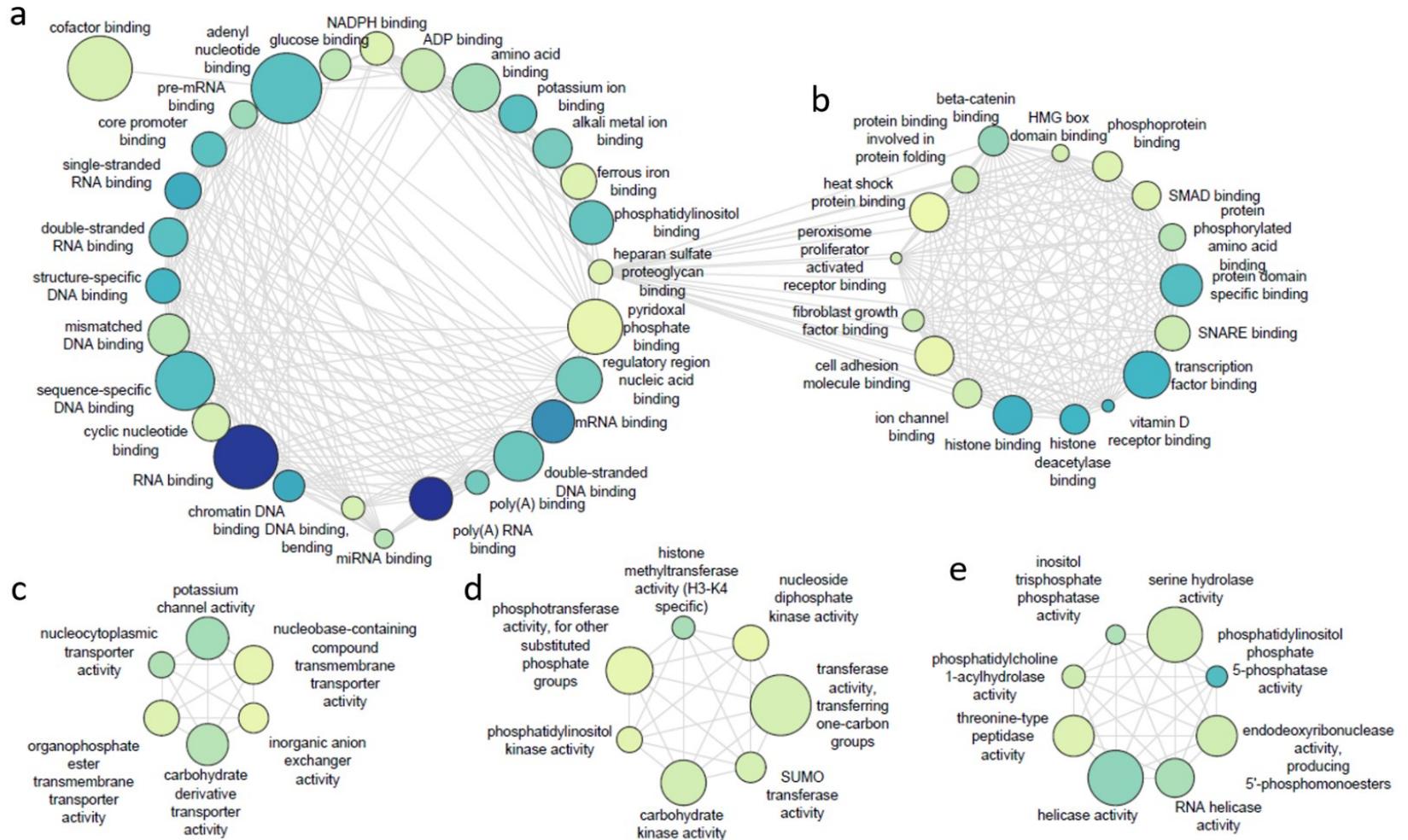


Figure 4.4 Enriched GO Molecular Function semantic relation network in heart of dehydrated *Xenopus laevis*. Enriched GO Molecular Function terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation network of summary terms was plotted using Cytoscape. The largest densely connected network comprised two sub-networks related to DNA and RNA binding (**a**) and transcription factor binding (**b**). Three small networks related to transporter and channel activities (**c**), post-translational modification activities (**d**), and molecular cleavage activities (**e**). Node size is proportional to the number of genes each term encompasses and the darker the blue color of a node is the lower the FDR-adjusted p -value and grey edges indicate a semantic relationship

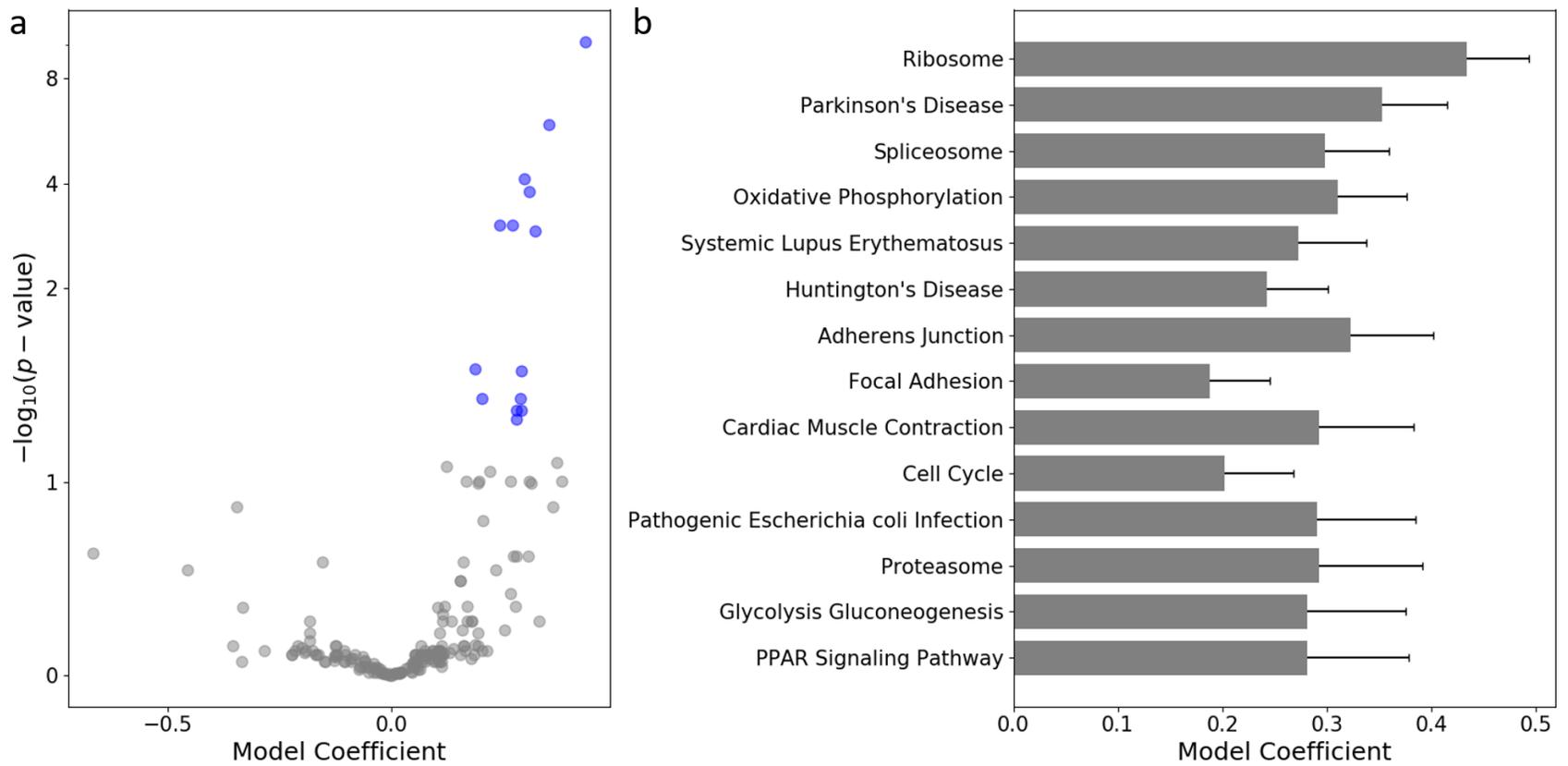


Figure 4.5 Enrichment of KEGG pathways in heart of dehydrated *Xenopus laevis*. **a)** Volcano plot of KEGG pathway enrichment. **b)** All significantly enriched KEGG pathways. All other information as in **Figure 4.3**

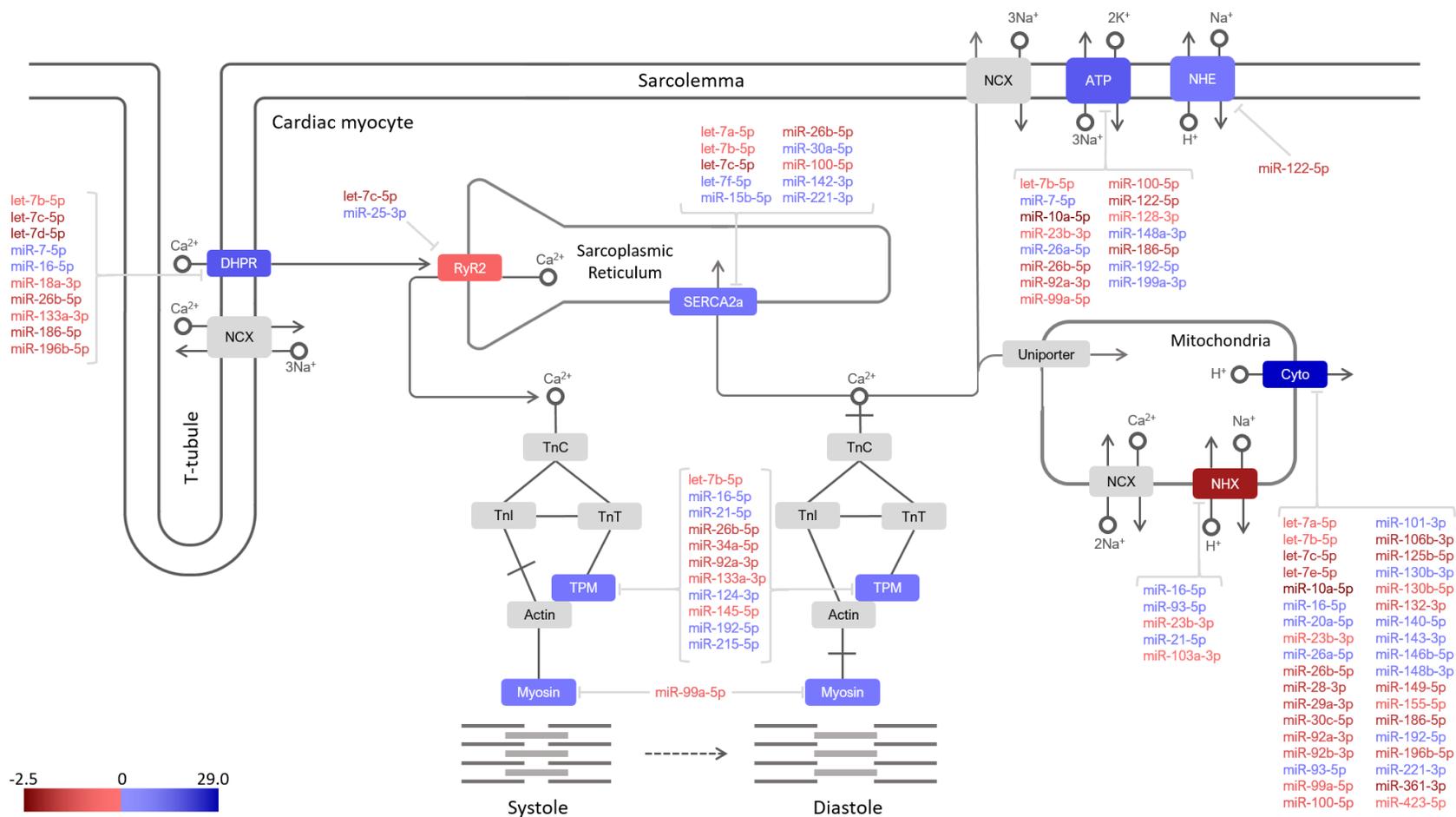


Figure 4.6 MicroRNA that target genes in the Cardiac Muscle Contraction KEGG pathway. Genes and microRNA are colored by their S_{mRNA} and $S_{microRNA}$ scores respectively, where a darker blue color indicates a more positive score and a darker red color indicates a more negative score. S_{mRNA} is equal to the negative of the summation of $S_{microRNA}$ scores of all microRNA that target that mRNA. For more information about S_{mRNA} and $S_{microRNA}$ scores see 4.5 Materials and Methods for definitions. A full list of genes encompassed by each label is found in **Supplementary Table 4.9**

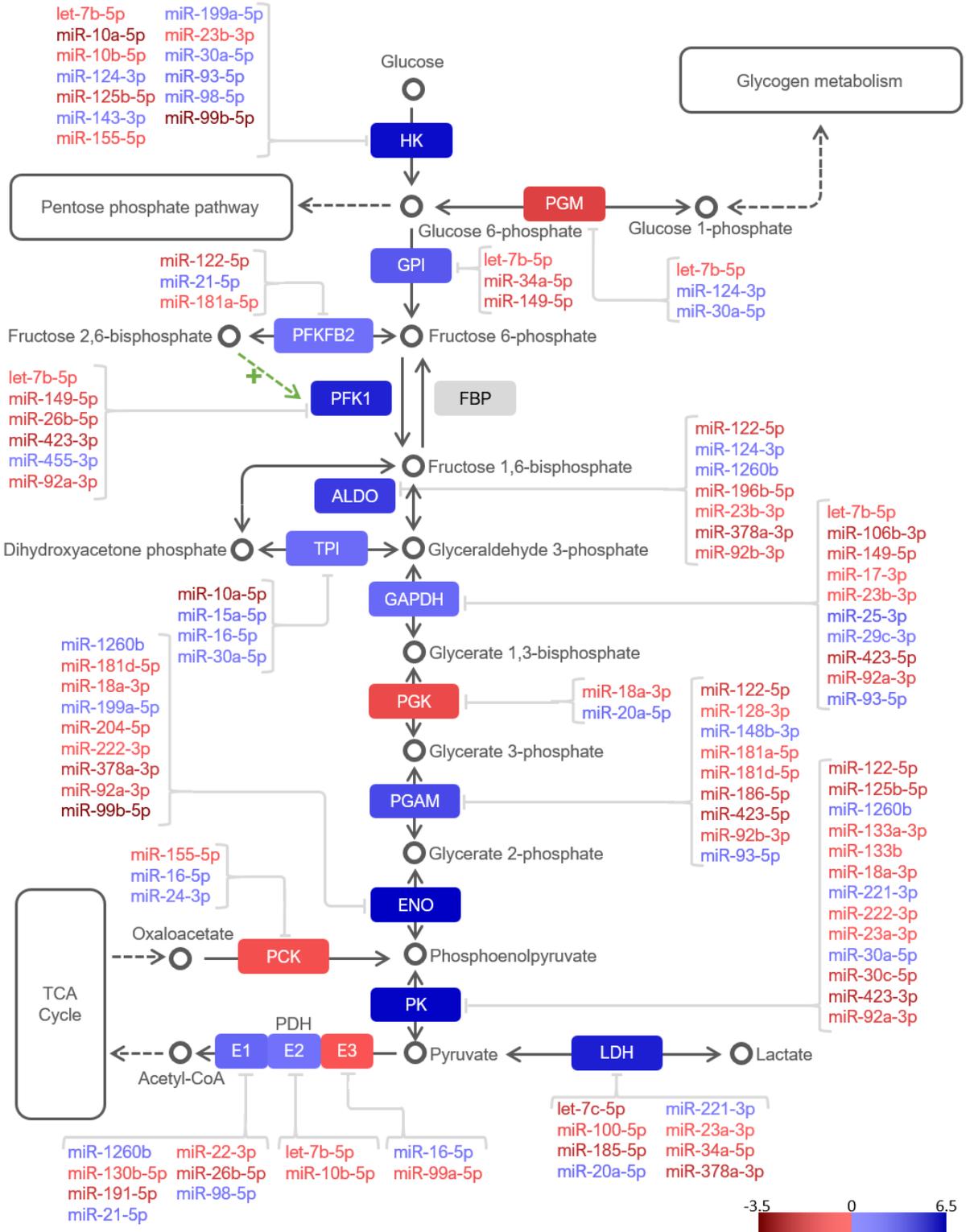


Figure 4.7 MicroRNA that target genes in the Glycolysis Gluconeogenesis KEGG pathway. All other information as in **Figure 4.6**. A full list of genes encompassed by each label is found in **Supplementary Table 4.10**

Supplementary Table 4.1 Primers used for RT-qPCR validation of microRNA-seq

MicroRNA	Sequence	Primer
let-7d-5p	AGAGGUAGUAGGUUGCAUAGUU	ACACTCCAGCTGGGAGAGGTAGTAGGTTGCATAGTT
mir-10a-5p	UACCCUGUAGAUCCGAAUUUGUG	ACACTCCAGCTGGGTACCCTGTAGATCCGAATTTGTG
mir-129-1-3p	AAGCCCUUACCCCAAAAAGUAU	ACACTCCAGCTGGGAAGCCCTTACCCCAA
mir-137-5p	ACGGGUAAUUCUUGGGUGGAUAAU	ACACTCCAGCTGGGACGGGTATTCTTGGGT
mir-184	UGGACGGAGAACUGAUAAAGGGU	ACACTCCAGCTGGGTGGACGGAGAACTGATAAGG
mir-24-1-5p	UGCCUACUGAGCUGAUUUCAGU	ACACTCCAGCTGGGTGCCTACTGAGCTGATATCA
mir-30a-3p	CUUUCAGUCGGAUGUUUGCAGC	ACACTCCAGCTGGGCTTTCAGTCGGATGTTTGCA
mir-3184-5p	UGAGGGGCCUCAGACCGAGCUUUU	ACACTCCAGCTGGGTGAGGGGCCTCAGAC
mir-378c	ACUGGACUUGGAGUCAGAAGAGUGG	ACACTCCAGCTGGGACTGGACTTGGAGTCAG
mir-423-3p	AGCUCGGUCUGAGGCCCCUCAGU	ACACTCCAGCTGGGAGCTCGGTCTGAGGC
mir-486-5p	UCCUGUACUGAGCUGCCCCGAG	ACACTCCAGCTGGGTCTGTACTGAGCTGC
Reference Genes		
mir-27a-3p	UUCACAGUGGCUAAGUUCCGC	ACACTCCAGCTGGGTTCACAGTGGCTAAGTTCC
mir-15b-5p	UAGCAGCACAUCAUGGUUUACA	ACACTCCAGCTGGGTAGCAGCACATCATGGT
mir-181b-3p	CUCACUGAACAAUGAAUGCAA	ACACTCCAGCTGGGCTCACTGAACAATGAATGCA
mir-103a-3p	AGCAGCAUUGUACAGGGCUAUGA	ACACTCCAGCTGGGAGCAGCATTGTACAGG
mir-103b	UCAUAGCCCUGUACAAUGCUGCU	ACACTCCAGCTGGGTTCATAGCCCTGTACAATGC
Universal Reverse Primer	-	CTCACAGTACGTTGGTATCCTTGTG

Supplementary Table 4.2 MicroRNA-Seq summary and raw read processing results

Sample	Total Reads	Reads with Adapters	Reads Passed Filter	Negative Filtered	Mature Aligned	Hairpin Aligned	Unaligned
Control 1	9,115,490	8,081,543	8,067,908	2,764,724	1,267,074	180,893	3,855,217
Control 2	7,439,618	6,759,080	6,834,313	2,085,041	1,818,752	276,145	2,654,375
Control 3	8,994,036	7,966,768	8,501,099	2,124,865	2,516,138	384,237	3,475,859
Dehydration 1	8,520,980	7,517,837	7,944,915	2,192,147	1,937,140	309,522	3,506,106
Dehydration 2	8,430,536	7,569,183	8,006,690	2,210,857	2,270,511	392,394	3,132,928
Dehydration 3	8,001,709	7,181,757	7,191,606	1,901,737	1,920,558	546,163	2,823,148
Dehydration 4	7,822,058	7,214,714	7,111,407	2,322,067	1,873,199	297,429	2,618,712
Total	58,324,427	52,290,882	53,657,938	15,601,438	13,603,372	2,386,783	22,066,345
	8,516,381 ±	7,602,464 ±	7,801,107 ±	2,324,877 ±	1,867,321 ±	280,425 ±	3,328,484 ±
Control	934,480	732,643	864,830	381,439	625,947	101,740	613,836
	8,193,821 ±	7,370,873 ±	7,563,655 ±	2,156,702 ±	2,000,352 ±	386,377 ±	3,020,224 ±
Dehydration	335,749	200,894	477,699	179,390	182,132	114,581	386,796

Supplementary Table 4.3 Differential expression results of microRNAs in heart of *Xenopus laevis* exposed to dehydration (Data available in: Hawkins 2020)

Supplementary Table 4.4 GO Molecular Functions enriched by microRNAs differentially expressed in the heart of *Xenopus laevis* during dehydration (Data available in: Hawkins 2020)

Supplementary Table 4.5 GO Biological Processes enriched by microRNAs differentially expressed in the heart of *Xenopus laevis* during dehydration (Data available in: Hawkins 2020)

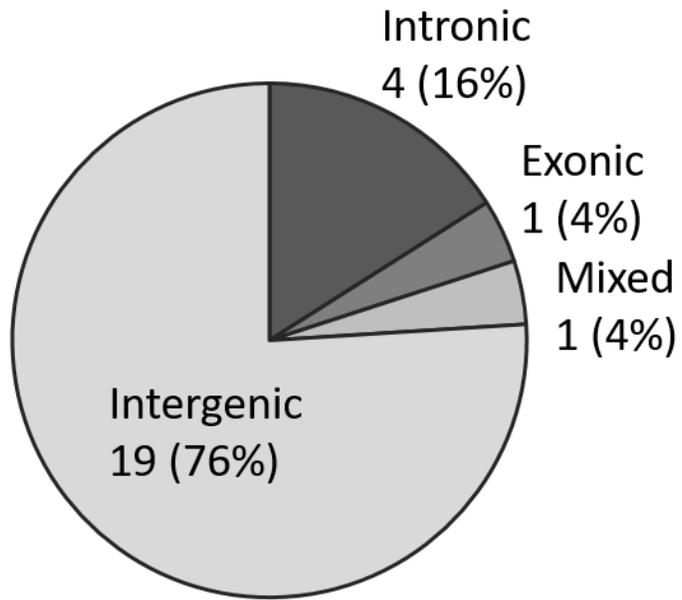
Supplementary Table 4.6 GO Cellular Components enriched by microRNAs differentially expressed in the heart of *Xenopus laevis* during dehydration (Data available in: Hawkins 2020)

Supplementary Table 4.7 KEGG pathways enriched by microRNAs differentially expressed in the heart of *Xenopus laevis* during dehydration (Data available in: Hawkins 2020)

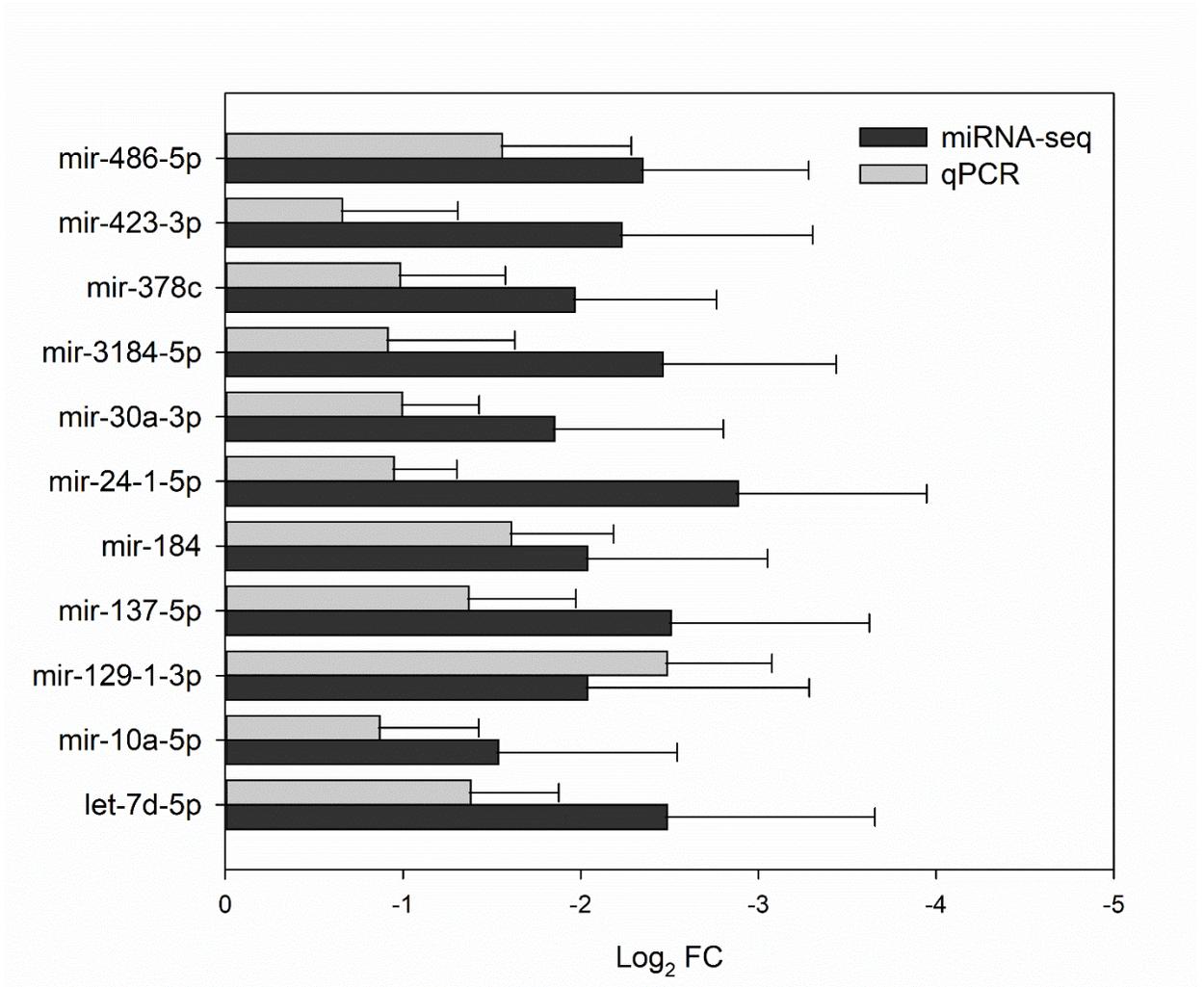
Supplementary Table 4.8 Predicted MicroRNA:mRNA target pairing of microRNAs expressed in the heart of *Xenopus laevis* in response to dehydration (Data available in: Hawkins 2020)

Supplementary Table 4.9 MicroRNAs expressed in the heart of *Xenopus laevis* during dehydration that target genes in the KEGG Cardiac Muscle Contraction pathway (Data available in: Hawkins 2020)

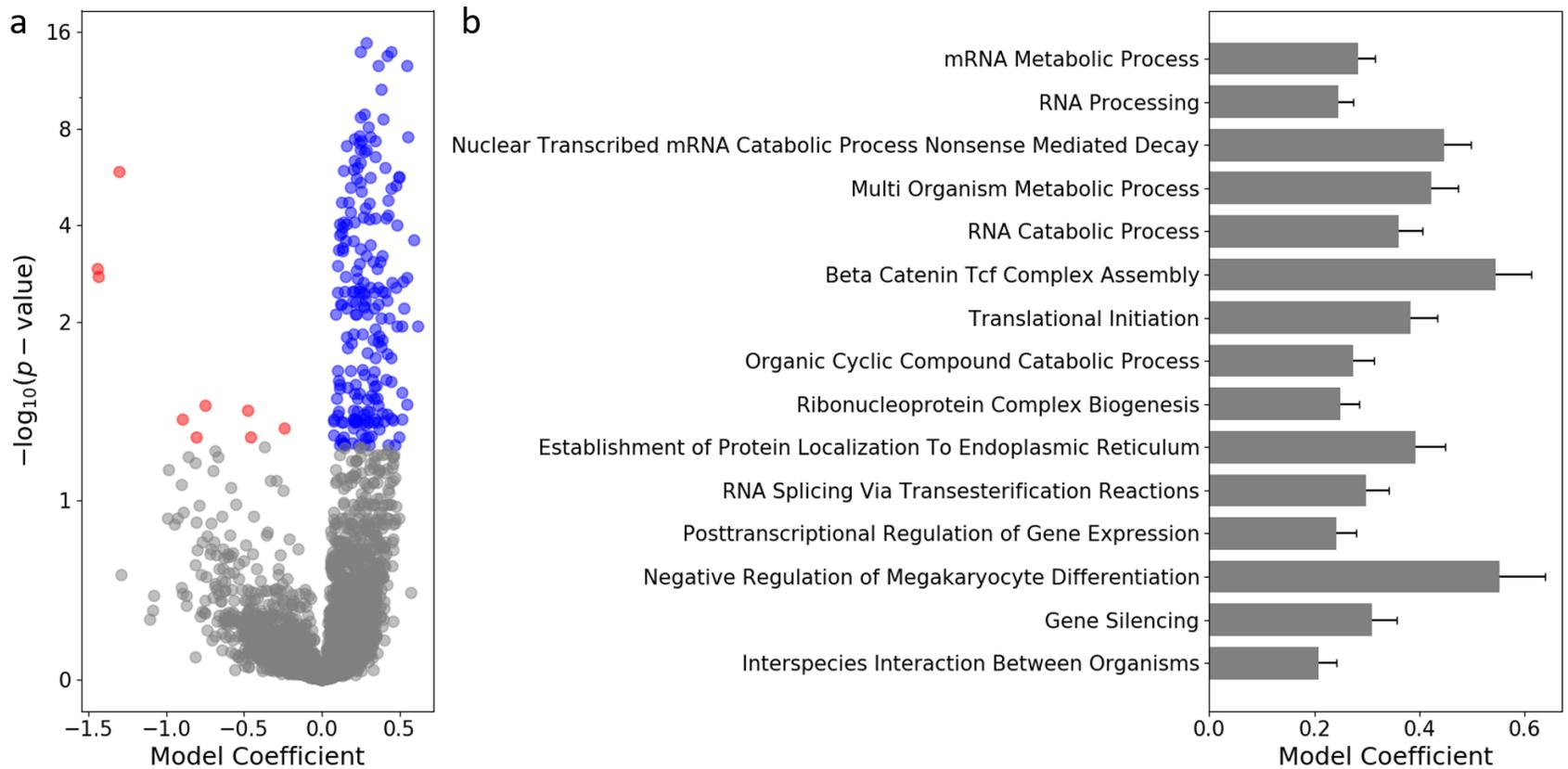
Supplementary Table 4.10 MicroRNAs expressed in the heart of *Xenopus laevis* during dehydration that target genes in the KEGG Glycolysis Gluconeogenesis pathway (Data available in: Hawkins 2020)



Supplementary Figure 4.1. Genomic location of significantly differentially expressed microRNAs. Mixed indicates the microRNA is located in the genome where multiple genes overlap.

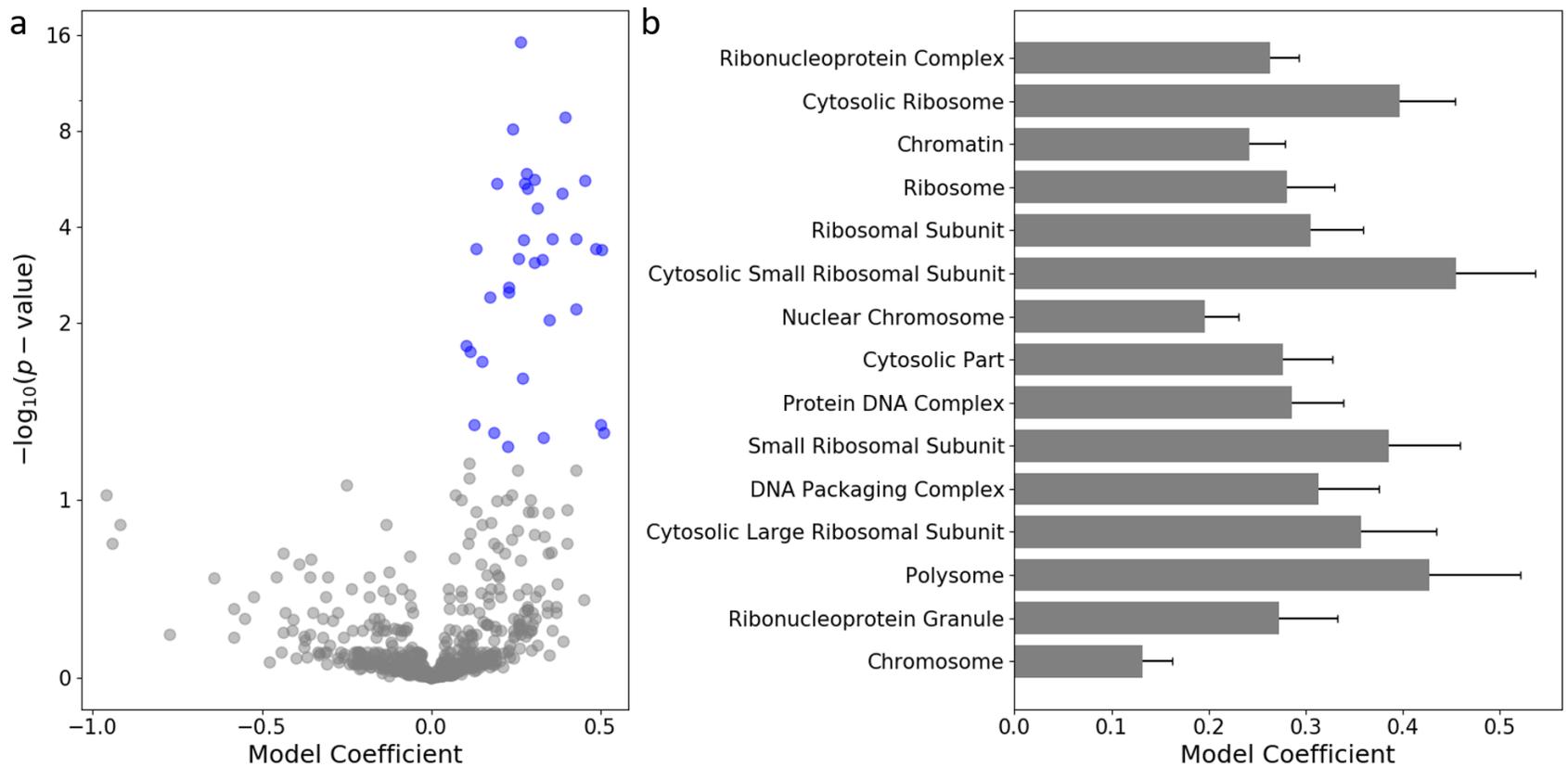


Supplementary Figure 4.2 RT-qPCR Validation results of differentially expressed microRNA



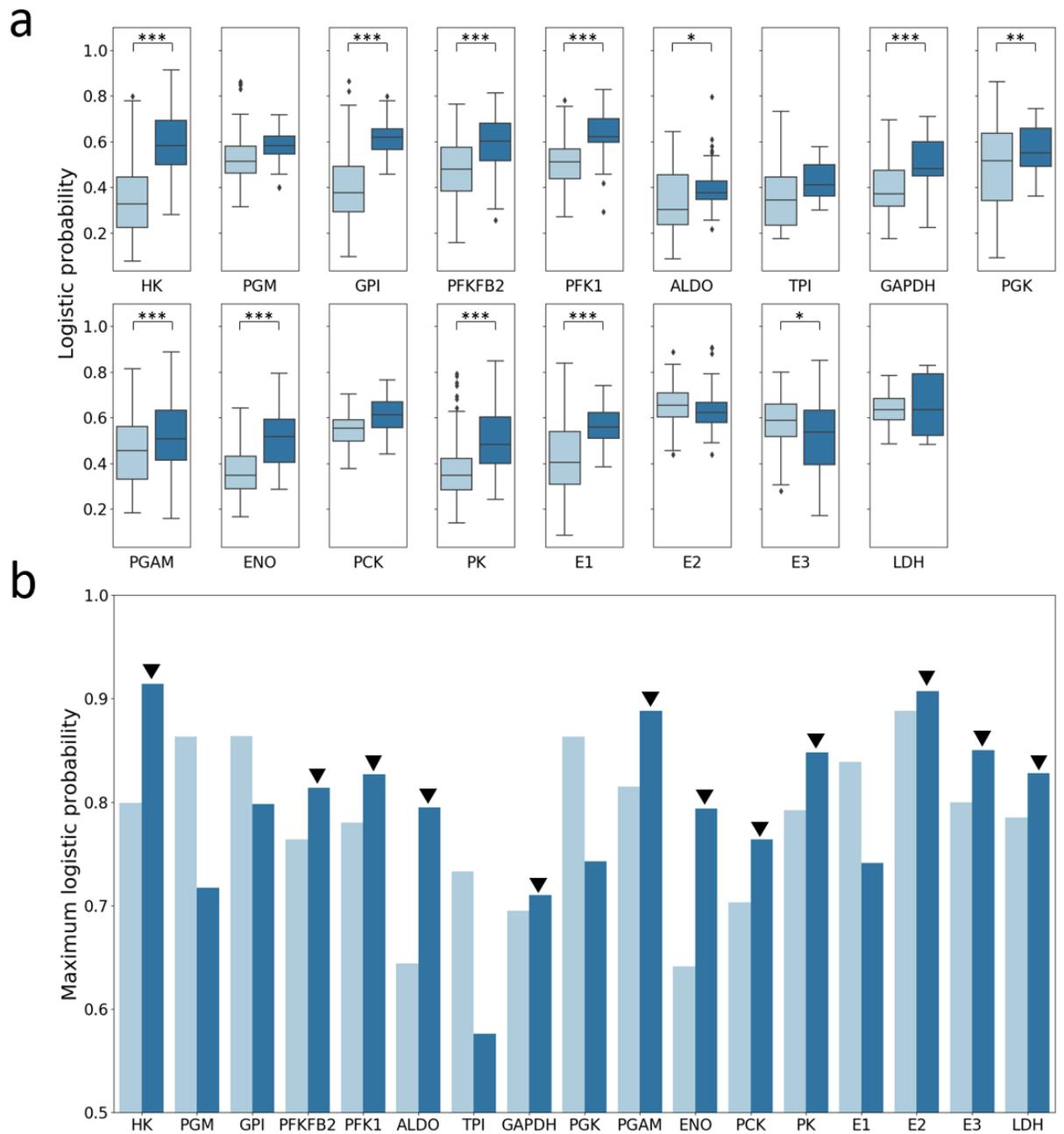
Supplementary Figure 4.3 Enrichment of GO Biological Processes terms in heart of dehydrated *Xenopus laevis*. **a)** Volcano plot of GO Biological Processes term enrichment with an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. **b)** Top 15 most significantly enriched (by FDR-adjusted p -value) GO Biological Processes terms

Supplementary Figure 4.4 Enriched GO Biological Processes semantic relation network in heart of dehydrated *Xenopus laevis*. Enriched GO Biological Processes terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation network of summary terms was plotted using Cytoscape



Supplementary Figure 4.5 Enrichment of GO Cellular Components terms in heart of dehydrated *Xenopus laevis*. **a)** Volcano plot of GO Cellular Components term enrichment with an FDR-adjusted p -value threshold of 0.05. Red and blue circles indicate significantly enriched GO terms with negative and positive model coefficients, respectively. Model coefficients are a predicted measure of microRNA regulation of that term in dehydrated animals relative to control animals. A positive coefficient indicates downregulation of microRNA targeting the genes encompassed in that term, thus potentially upregulating the genes and a negative coefficient indicates the opposite. **b)** Top 15 most significantly enriched (by FDR-adjusted p -value) GO Cellular Components terms

Supplementary Figure 4.6 Enriched GO Cellular Components semantic relation network in heart of dehydrated *Xenopus laevis*. Enriched GO Cellular Components terms were summarized by REVIGO (Supek et al. 2011) and the resulting semantic relation network of summary terms was plotted using Cytoscape



Supplementary Figure 4.7 Comparison human and *Xenopus laevis* microRNAs targeting of 3'UTRs of glycolytic genes. **a**) STarMir (<http://sfold.wadsworth.org/cgi-bin/starmirtest2.pl>) was used to determine logistic probabilities of microRNAs targeting 3'UTRs of human and *X. laevis* genes involved in glycolysis and gluconeogenesis from **Figure 4.7**. Statistical differences were determined with Student's t-test where * indicates p -value < 0.05, ** p -value < 0.01, and *** p -value < 0.001. **b**) Maximum logistic probability for each gene. Black arrows indicate genes where *X. laevis* had a higher maximum logistic probability than humans

Chapter 5

Phosphoproteomic analysis of *Xenopus laevis* reveals expression and phosphorylation of hypoxia inducible PFKFB3 during dehydration

Liam J. Hawkins¹; Xiaoshuang Wang²; Xiaomin Xue²; Hui Wang^{2,*}; Kenneth B. Storey^{1,*}

¹ Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada

² Key Laboratory of Animal Physiology, Biochemistry and Molecular Biology of Hebei Province, College of Life Sciences, Hebei Normal University, Shijiazhuang, Hebei 050024, China

5.1 Abstract

The African clawed frog, *Xenopus laevis*, tolerates severe dehydration when their aquatic environments evaporate during arid summer months. Dehydration tolerance in this species is predicated on multiple tissue specific adaptations including switching from ammonotelism to ureotelism and from aerobic to anaerobic metabolism. Reversible protein phosphorylation is a mechanism for enacting these adaptations, therefore we performed phosphoproteomic analysis on the liver and skeletal muscle of *X. laevis* exposed to severe dehydration to investigate potential adaptations. The liver showed strong changes in phosphorylation patterns compared to skeletal muscle. These liver specific changes were enriched in cellular functions and pathways related to central metabolic pathways including glycolysis/gluconeogenesis, the TCA cycle, and amino acid metabolism, whereas both tissues were enriched for structural and cytoskeletal elements. The prominence of phosphoproteins related to glucose metabolism seen here led us to find that the hypoxia inducible PFKFB3 enzyme was highly phosphorylated at S463 (homologous to human S461) and therefore activated during dehydration. We then examined expression of the four transcript variants of the *pfkfb3* gene and found all to be upregulated during dehydration, which may be due to the eight hypoxia responsive elements in the *pfkfb3* promoter. This trend was not found with the liver PFKFB isozyme, PFKFB1, which although potentially activated by dephosphorylation, was downregulated during dehydration. These results further support the role of anaerobic glycolysis during dehydration in *X. laevis* and elucidate a potential mechanism for its increased rate.

5.2 Introduction

The African clawed frog (*Xenopus laevis*) is one of a select group of amphibians that encounters and tolerates severe dehydration in their native habitats. When the bodies of water they occupy dry up during the summer months, these mostly aquatic animals are exposed to evaporative dehydration and can survive beyond 30% body water loss (Hillman 1978a). Interestingly, while *X. laevis* have been intensely characterized within the context of their role as model organisms for developmental and cellular biology, their adaptations necessary for surviving dehydration are understudied.

It is known that the response to dehydration in *X. laevis* manifests in a tissue specific manner. At an organismal level, blood is preferentially distributed during dehydration such that skeletal muscle has the greatest loss while blood distribution to the brain actually increases (Hillman and Sommerfeldt 1981). This results in disparate levels of dehydration stress that individual tissues have adapted to by harbouring different functions in response to dehydration. For example, the skin secretes a mucus layer to minimize evaporative dehydration, the heart sustains a high rate to maximize oxygen delivery (Hillman 1978a), and the liver switches from ammonotelism to ureotelism to avoid a toxic build up nitrogenous waste (Balinsky et al. 1961). At high levels of dehydration, *X. laevis* increase rates of anaerobic glycolysis as apparent by elevated whole-body lactate levels (Hillman 1978a).

Investigation into metabolic adaptations have focused on glucose metabolism due to the dependency on anaerobic glycolysis, particularly in the liver. In this tissue, enzymology studies have indicated that pyruvate kinase (PK) has a greater activity during dehydration than under control conditions (Dawson et al. 2018). This is of importance since it is an ATP producing step in glycolysis and generally considered to be a rate-

limiting enzyme under tight regulation in the liver. Increased activity of PK in the liver would support the increased rate of anaerobic glycolysis which is further evident by a study examining lactate dehydrogenase (LDH). LDH in this animal was shown to have decreased affinity for its substrates, although this affinity is restored in the presence of urea concentrations seen during dehydration (Katzenback et al. 2014). In skeletal muscle, hexokinase also has altered kinetics (Childers and Storey 2016) that may regulate glucose entry into muscle cells when plasma glucose levels are elevated during dehydration (Malik and Storey 2009).

The commonality between the above-mentioned glycolytic enzyme studies is the observed changes in phosphorylation state that provide a mechanism to explain their altered kinetics during dehydration. Phosphorylation is a quick and reversible method of altering protein function and is a known regulator of metabolic pathways and cellular processes during environmental stress responses (Storey and Storey 1990; Storey 2002; Cowan and Storey 2003). By examining changes in phosphorylation patterns, we can determine which pathways and processes are being actively regulated in response to a variety of environmental stresses and broaden our understanding of adaptations necessary for tolerating them. This has been explored in a variety of species including hibernating mammals (Chung et al. 2013; Vertommen et al. 2017) and most recently in the freeze-, anoxia- and dehydration tolerant frog species, *Rana sylvatica* (Hawkins et al. 2019).

Like *X. laevis*, *R. sylvatica* is tolerant to extreme levels of body water loss (Churchill and Storey 1994). Phosphoproteomic analysis indicates that several enzymes involved in glycogenolysis and glycolysis are differentially phosphorylated, and may function to mobilize glycogen stores to supply fuel for anaerobic glycolysis during dehydration

(Hawkins et al. 2019). This analysis proved fruitful for expanding what we know about how this anuran can tolerate multiple stresses.

Here we take a similar approach to analyze changes in the phosphoproteome of *X. laevis* in response to dehydration exposure. We exposed adult *X. laevis* to dehydrating conditions and used LC-MS/MS to quantify phosphopeptides from the liver and skeletal muscle of these animals. This is followed by bioinformatic analysis to gain an overview of the processes and pathways responding to dehydration and broaden our understanding of dehydration tolerance.

5.3 Materials and Methods

5.3.1 Animal experiments

Adult male *X. laevis* from the University of Alberta Science Animal Support Services were shipped to Carleton University where they were acclimated for one week at 22°C in aerated tanks filled with dechloraminated water. *X. laevis* were fed three times per week and the tank water was changed each feeding day. The animals were then randomly divided into control and dehydration exposure groups (mean body mass 64.40 ± 10.77 g and 64.58 ± 10.78 g respectively). Control animals were maintained in the above conditions. Animals in the dehydration exposure group were placed in dry tanks at 22°C and weighed periodically over the course of two days until they had lost approximately 30% body water (actual % body water loss 33.59 ± 2.58). Percent body water loss was calculated using the following equation:

$$\% \text{ Body water loss} = \frac{m_i - m_d}{m_i \times BWC_i} \times 100\%$$

where m_i is the initial mass of the animal, m_d is the dehydrated mass, and BWC_i is initial body water content as previously determined (Malik and Storey 2009). Control and dehydration exposed animals were then euthanized by pithing and skeletal muscle and liver tissues were quickly dissected, frozen in liquid nitrogen, and stored at -80°C until use. All animal protocols were approved by the Carleton University Animal Care Committee (protocol #106936) and conformed with the guidelines of the Canadian Council on Animal Care.

5.3.2 Protein extraction

Frozen skeletal muscle and liver tissue samples weighing 200 mg (n = 3 from each experimental condition) were extracted separately using EMD Millipore Cell Lysis Buffer (catalogue #43-040) with 1 mM EDTA, 1 mM EGTA, phosphatase inhibitors (1 mM Na₃VO₄, 10 mM NaF, 10 mM β-glycerophosphate) and protease inhibitors (BioShop catalogue #PIC001). Tissues were homogenized with a Dounce homogenizer in 1:4 w:v lysis buffer and incubated for 30 min on ice with intermittent vortexing. Total soluble protein lysates were then collected by centrifuging samples at 14,000 x g for 20 min at 4°C and taking the supernatant. Protein concentrations were determined using Bradford assay (Bio-Rad Cat #500-0005), then samples were frozen and lyophilized for 12 h using a Labconco FreeZone 12L freeze dryer and stored at room temperature until use.

5.3.3 Phosphoproteomic analysis

Phosphoproteomic analysis of samples was performed as previously described (Hawkins et al. 2019). Briefly, samples were trypsin digested then enriched for phosphopeptides using TiO₂ beads. Enriched phosphopeptides were analyzed using a nanoACQUITY UPLC M-Class system (Waters, USA) and an electrospray ionization Q Exactive HF mass spectrometer (Thermo Fisher Scientific, USA). Mass spectrometry results were analyzed using Proteome Discoverer 2.2 (Thermo Fisher Scientific, USA) with a database created from *X. laevis* protein sequences (56748 sequences, 2018/04/01) retrieved from UniProt.

5.3.4 Data analysis and differential phosphorylation analysis

For each tissue, non-phosphorylated peptides and phosphopeptides not present in at least two control and two dehydrated biological replicates were filtered out of the dataset.

For each phosphopeptide, \log_2 -fold changes and standard deviations between control and dehydrated animals were calculated using the following equations:

$$\log_2 FC = \log_2 \frac{AVG\left(\begin{bmatrix} Deh_1 \\ \vdots \\ Deh_n \end{bmatrix}\right)}{AVG\left(\begin{bmatrix} Con_1 \\ \vdots \\ Con_n \end{bmatrix}\right)}$$

$$\log_2 SD = \sqrt{SD\left(\log_2 \begin{bmatrix} Con_1 \\ \vdots \\ Con_n \end{bmatrix}\right)^2 + SD\left(\log_2 \begin{bmatrix} Deh_1 \\ \vdots \\ Deh_n \end{bmatrix}\right)^2}$$

Student's t -test with a Benjamini-Hochberg multiple test correction was used to determine statistical significance (FDR corrected p -value < 0.05).

5.3.5 Hierarchical clustering and principle component analysis

Hierarchical clustering was performed on samples using the *seaborn* python package (Waskom 2018) which uses the *scipy* python package (Jones et al. 2001). Samples were clustered based on the Euclidean distance between the relative quantities of their respective phosphopeptides and phosphopeptides were clustered based on the Euclidean distance between their relative quantity in each sample. Principle component analysis (PCA) of samples was performed using the *sklearn* python package (Pedregosa et al. 2011).

5.3.6 Gene set analysis

Overrepresentation analysis of gene ontology (GO) terms (The Gene Ontology Consortium 2019) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa et al. 2019) was performed using the WebGestalt webserver (Liao et al. 2019). While WebGestalt does not have built-in support for *X. laevis*, custom GO and KEGG

annotation databases can be created and used with this tool. First GO and KEGG annotations for all *X. laevis* proteins were retrieved from UniProt (UniProt Consortium 2019) and GO and KEGG gene matrix transposed (GMT) files were created using a custom python script to be used as the functional database in WebGestalt. The proteins corresponding to significantly differentially expressed (FDR-corrected p -value < 0.05) phosphopeptides were uploaded as the gene list, all proteins with GO or KEGG annotations were uploaded as the background reference list of genes and the Benjamini-Hochberg procedure was used as a multiple test correction.

Over represented GO terms were further explored using REVIGO (Supek et al. 2011) which produces network diagrams of semantic similarity between GO terms using the hierarchical graph structure of GO. Significantly enriched GO terms and their FDR-corrected p -values produced by WebGestalt were uploaded using the default REVIGO settings and network diagrams of these significantly enriched terms were retrieved.

5.3.7 RNA extraction and cDNA synthesis

For RNA extraction, 50 mg of tissue from each sample was homogenized using a Polytron homogenizer in 1 mL of TRIzol reagent (Invitrogen; Cat# 15596-018). A 200 μ L volume of chloroform was added to each sample which was then centrifuged at $10,000 \times g$ at 4°C for 15 min. The aqueous phase containing total RNA was removed and precipitated using 500 μ L of 2-propanol for 10 min at room temperature, then pelleted at $10,000 \times g$ at 4°C for 15 min. Each pellet was then washed twice with ethanol and air-dried before resuspension in 50 μ L of RNase-free water. RNA concentration and purity were determined using a BioTek Take3 microspot plate and PowerWave HT microplate

spectrophotometer where all samples had 260/280 ratios of ~2.0, then RNA integrity was verified using a 1% agarose gel.

Reverse transcription cDNA synthesis was performed on 2 µg of RNA from each sample in 10 µL of RNase-free water using the Invitrogen M-MLV Reverse Transcriptase Kit (Cat# 28025013). A 1 µL volume of 200 ng·µL⁻¹ oligo(dT) (Sigma-Aldrich) was added to each sample and incubated at 65°C for 5 min. Samples were then cooled on ice and the reverse transcriptase reaction was performed according to the manufacturer's directions using 1 µL of 10 mM dNTPs (BioShop, Canada).

5.3.8 Primer design and RT-qPCR

Primers were designed using Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) using default settings with the exception of a PCR product size of 80-120 bp, and primer pair specificity checking for *Xenopus laevis* (taxid: 8355). Primers were synthesized by Integrated DNA Technologies and sequences are shown in **Supplementary Table 5.1**.

RT-qPCR was performed using a Bio-Rad CFX Connect as previously described (Pellissier et al. 2006). PCR reactions were comprised of 8 µL cDNA, 2 µL qPCR buffer (100 mM Tris-HCl [pH 8.5], 500 mM KCl, 1.5% v:v Triton X-100, and 20 mM MgCl₂), 0.16 µL of 25 mM dNTPs, 4 µL of 1 M trehalose, 0.5 µL of formamide, 0.025 µL of 100X SYBR Green in DMSO (Invitrogen; #S7585), 0.5 µL of 0.3 nmol·µL⁻¹ of each primer, 0.125 µL of 5U·µL⁻¹ wild-type Taq (BioShop; #TAQ001.1) and autoclaved ddH₂O for a final volume of 20 µL. The following thermocycle series was used: An initial 95°C denaturing step for 3 min, then 40 cycles of denaturing at 95°C for 10 sec, annealing at 57°C for 20 sec, and extension at 72°C for 20 sec. A melting analysis was

performed from 55-95°C to ensure the amplification of single products, and a 2-fold dilution standard curve for each target was run on quantification plates to ensure reaction efficiencies between 90-110%. Suitable reference genes in liver tissue were previously determined to be *atp5f1b.L* and *sdha.L* (Hawkins et al. 2018) and quantification was performed using Bio-Rad CFX Maestro 1.1 which employs the Pfaffl method (Pfaffl 2001).

5.3.9 Data visualization

Unless otherwise stated figures were created using the *matplotlib* python package (Hunter 2007) and SigmaPlot 12.5. Hierarchical clustering was visualized using the *seaborn* python package (Waskom 2018) and visualization of REVIGO GO network diagram was done using Cytoscape (Su et al. 2014). *De novo* protein modeling of *X. laevis* PFKFB1 and PFKFB3 was done using I-TASSER (Yang et al. 2015) and visualized using PyMOL (<https://pymol.org>).

5.4 Results

5.4.1 Phosphoproteomics summary

Phosphoproteomic analysis resulted in the identification of 787 unique phosphopeptides in the liver and 692 in skeletal muscle, which matched to 824 and 396 proteins respectively. Both liver and skeletal muscle showed similar proportions of phosphorylated residues (**Supplementary Figure 5.1a**). In the liver, there were 715 (82.66%) phosphoserine residues, 130 (15.03%) phosphothreonine residues, and 20 (2.31%) phosphotyrosine residues. In skeletal muscle, there were 633 (80.74%) phosphoserine, 134 (17.09%) phosphothreonine and 17 (2.17%) phosphotyrosine residues. The number of phosphorylated residues per phosphopeptide was also comparable between tissues (**Supplementary Figure 5.1b**). In the liver, 711 (90.34%) phosphopeptides had one phosphoresidue, 74 (9.40%) had two phosphoresidues and 2 (0.25%) had three or more phosphoresidues. In muscle, 608 (87.86%) phosphopeptides had one phosphoresidue, 76 (10.98%) had two phosphoresidues, and 8 (1.16%) had three or more phosphoresidues.

5.4.2 Dehydration produces distinct phosphorylation patterns

Hierarchical clustering of relative quantities of all phosphopeptides resulted in clustering of control samples together and dehydration samples together in liver (**Figure 5.1a**). This contrasted clustering in muscle which showed intermixing of control and dehydration samples when using all phosphopeptides (**Figure 5.1b**), indicated less of a defined phosphorylation pattern differentiating control and dehydrated animals in this tissue. Unsurprisingly, when hierarchical clustering was performed on only statistically

differentially phosphorylated peptides both tissues showed clear clustering of samples from control and dehydrated animals (**Supplementary Figure 5.2**).

PCA resulted in similar conclusions to hierarchical clustering results. In liver, samples from control and dehydrated animals clustered together, and were particularly separable along the first component (explaining 69% of the variance between samples) (**Figure 5.2a**), whereas there was less separation and clustering of control and dehydrated muscle samples (**Figure 5.2b**). These results again support a stronger distinct phosphorylation pattern in the liver than in muscle. When analyzed together, PCA analysis showed higher inter-tissue separation than separation between experimental conditions (**Figure 5.2c**), although this is expected given the dissimilarity between liver and muscle tissue. It should be noted that when PCA analysis was performed on both tissues together, only the subset of phosphopeptides that appeared in both tissues could be examined (125 phosphopeptides), versus a much higher number that appeared in each tissue individually (787 in liver, 692 in muscle). When PCA was performed on tissues separately and together, cumulative explained variance curves were similar (**Figure 5.2d**).

5.4.3 Differential phosphorylation in response to dehydration

Significant differences in relative phosphopeptide amounts varied between the two tissues to a much greater extent than would be predicted by the number of phosphopeptides found in each tissue (**Figure 5.3**). In the liver, 308 (39.14%) phosphopeptides had FDR-corrected p -values < 0.05 (**Figure 5.3a**), while in muscle only 8 (1.2%) phosphopeptides met this criteria (**Figure 5.3b**). This discrepancy appears to be due to intra-experimental condition variability since the number and distribution of phosphopeptides with high and low fold-changes is comparable between liver and muscle

(**Figure 5.3b**). These results further support the distinct phosphorylation patterns present in the liver and not muscle as seen in the clustering analysis.

By comparing proteins with corresponding phosphopeptides found in each tissue, the overlap of proteins is substantial where 181 proteins are found in common (21.97% of liver proteins, 45.71% of muscle proteins) (**Figure 5.4a**). When only proteins with corresponding significantly differentially abundant phosphopeptides are considered this number drops to 6 proteins (1.51% of liver proteins, 40.00% of muscle proteins) (**Figure 5.4b**), indicating that the vast majority of affected proteins in the liver are liver-specific, whereas the majority of affected proteins in the muscle are not muscle-specific but are also affected in the liver.

5.4.4 Overrepresentation of GO terms and KEGG pathways

REVIGO analysis revealed enrichment of differentially abundant phosphopeptides in small networks of semantically related non-redundant GO biological processes in the liver (**Figure 5.5**). These include a glucose metabolism related network (glycolytic process, fructose metabolism, citrate metabolism, fructose 2,6-bisphosphate metabolism, and protein autophosphorylation), a post-transcriptional regulation network (negative regulation of translational initiation, negative regulation of I κ B kinase/NF κ B signaling, and intracellular signal transduction), and a RNA processing network (RNA 3'-end processing and mRNA processing), along with other unconnected nodes (protein transport, plasma membrane tubulation, cell adhesion, actin filament organization, and RNA polyadenylation). GO cellular compartments were also enriched, namely microtubule and cytoskeletal compartments (**Supplementary Figure 5.3**), while enriched

GO molecular functions included an actin related network and a kinase related network (**Supplementary Figure 5.4**).

REVIGO analysis of muscle samples showed few enriched GO terms given the low number of differentially abundant phosphopeptides (**Figure 5.6**). Only a single GO biological process was enriched, regulation of striated muscle contraction (**Figure 5.6a**), two GO cellular compartments, cytoplasm and microtubule (**Figure 5.6b**), and two GO molecular functions, cysteine-type endopeptidase inhibitor activity and microtubule binding (**Figure 5.6c**).

KEGG pathways enriched for proteins with differentially abundant phosphopeptides in the liver pertained mainly to central metabolic pathways: fructose and mannose metabolism, pentose phosphate pathway, biosynthesis of amino acids, and glycolysis/gluconeogenesis, but also adherens junction and melanogenesis (**Figure 5.7**). No KEGG pathways were significantly enriched in the muscle.

5.4.5 Phosphorylation and expression of PFKFB enzymes

Seeing that central carbon pathways were enriched for proteins corresponding to differentially abundant phosphopeptides in the liver, we investigated which enzymes were contributing to this enrichment (**Table 5.1**). Two enzymes related to glycogen metabolism, glycogen synthase kinase-3 beta (GSK3 β) and phosphoglucomutase (PGM) had increased phosphorylation, whereas only one TCA cycle enzyme, ATP-citrate synthase (CS), showed significant differences in phosphorylation. Enzymes comprising glycolysis and gluconeogenesis, 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 1 and 3 (PFKFB1 and PFKFB3), the liver isoform of fructose-bisphosphate aldolase (ALDOB), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), did however all

show large changes in phosphorylation both in magnitude and statistical significance (**Table 5.1**).

Of particular interest is the phosphorylation pattern of PFKFB1 and 3 in the liver, the bifunctional enzymes that catalyze the interconversion of fructose 6-phosphate (Fru-6-P) and fructose 2,6-bisphosphate (Fru-2,6-P₂), a potent phosphofructokinase 1 (PFK1) activator (**Figure 5.8a**). PFKFB1 is the liver isoform and showed significantly decreased amount of phosphorylation at S31 and S32 (homologous to S33 and S34 of human PFKFB1). S31 is a known protein kinase A target within the phosphofructokinase-2 (PFK2) domain (**Figure 5.8b**), and when phosphorylated it suppresses PFK2 activity and activates fructose-2,6-bisphosphatase (FBP2) activity (Ros and Schulze 2013).

Analogous results are seen with the increased phosphorylation of PFKFB3 at S463 (homologous to S461 of human PFKFB3) (**Figure 5.8a**), which is part of the FBP2 domain that promotes PFK2 activity when phosphorylated (Bando et al. 2005) (**Figure 5.8c**).

Expression of PFKFB3 is not typical of liver tissue, however this isoform is also known as inducible PFKFB, as it has been observed to be expressed in response to hypoxic conditions of cancer cells by hypoxia inducible factor 1-alpha (HIF1 α) (Bando et al. 2005; Shi et al. 2017). This isoform has four transcript variants in *X. laevis* (**Figure 5.9a**), all of which code for proteins containing the regulatory S463 residue. We designed primers to measure the expression of variants 2, 3, and 4 by RT-qPCR, as well as a primer pair that would amplify all four variants (*pfkfb3 all*), and therefore give some measure incorporating variant 1 expression levels which otherwise has 100% identity with variants 2 and 4. Expression of variants 2, 3, and 4 increased 9.88 ± 1.43 , $4.44 \pm$

0.91, and 9.11 ± 1.22 -fold during dehydration, while the primer pair measuring all four variants increased 8.02 ± 1.00 -fold relative to control samples (**Figure 5.9b**). Increased expression of PFKFB enzymes is not a general response as seen by relative expression levels of PFKFB1 which actually decreased during dehydration (**Figure 5.9b**). These results may be explained by the presence of eight hypoxia-responsive elements (HREs, RCGTG (Mole et al. 2009)) within 2 kbp of the *pfkfb3* transcription start site (TSS) (five within ~500 bp) whereas only 2 are observed 2 kbp upstream of the *pfkfb1* TSS (**Figure 5.9c**).

5.5 Discussion

5.5.1 Consistent phosphorylation pattern in liver but not skeletal muscle

Our analysis of dehydration exposure in *X. laevis* identified a distinct dehydration-responsive phosphorylation pattern in liver, but not in skeletal muscle. These results are consistent with this animal's strategy of dehydration tolerance. In response to dehydration, the liver undergoes relatively large changes to its physiology and metabolism including converting from ammonotelism to ureotelism (Balinsky et al. 1967) and switching to anaerobic glycolysis as oxygen transport is hampered (Katzenback et al. 2014). Both processes are under the control of enzymes that can be phosphorylated to control their activity and kinetics, and thus our detection of dehydration-responsive phosphorylation patterns in this organ was expected. This is reflected in our clustering analysis where samples from control animals clustered separately from dehydration animals (**Figure 5.1a**). Likewise, distinct clustering of samples from control and dehydration animals was present in our PCA, where there was no overlap of confidence ellipses representing three standard deviations of the underlying Gaussian distribution (**Figure 5.2a**). In contrast to the liver, since these animals assume a water conserving posture to avoid unnecessary energy expenditure and evaporative dehydration, the muscles are dormant until rehydration occurs. This results in a less defined phosphorylation pattern in muscle relative to liver, where samples did not cluster with their experimental conditions (**Figure 5.1b**). Similarly, PCA of muscle samples showed much greater intra-condition variability and overlap of underlying distributions (**Figure 5.2b**). The overall difference between the tissues and their response to dehydration is seen when PCA was performed on all samples together. While this was

limited to the small subset of phosphopeptides present in both tissues, it is clear that there is greater differences at this level between tissues, than between experimental conditions within each tissue (**Figure 5.2c**). Again, this difference between tissues is unsurprising given the metabolic function and physiological role of these tissues in dehydration tolerance.

Contrasting results for significantly differentially abundant phosphopeptides in the two tissues supports a distinct liver response and higher variability in muscle. Liver showed a greater number of differentially abundant phosphopeptides (**Figure 5.3a,b**), the lack of which in muscle appears to be due to variation between muscle samples rather than a lack of identified phosphopeptides or distribution of average abundance within conditions (**Figure 5.3c**). Interestingly, while there is significant overlap between all proteins identified with corresponding phosphopeptides in liver and muscle (**Figure 5.4a**), when we look at only proteins with significantly differentially abundant phosphopeptides the overlap decreases below what would otherwise be expected (**Figure 5.4b**). Rather than maintaining the same ratio of overlapping proteins, the overlap drops to where most liver proteins are not differentially abundant in the muscle, but 6 of 15 proteins in the muscle are also differentially abundant in the liver. This further suggests that not only is the liver displaying a liver-specific phosphorylation pattern in response to dehydration, but in so far as the muscle is mounting a response, it appears unlikely as a substantial subset of changing phosphopeptides in muscle also change in liver. Possibly suggesting that these shared phosphopeptides represent a more global response to dehydration.

5.5.2 Central metabolic pathways are differentially phosphorylated

Functional analysis revealed enrichment of different processes in the liver and muscle. In the liver, central metabolism pathways are enriched for phosphoproteins including glycolytic/gluconeogenic pathways, the pentose phosphate cycle, biosynthesis of amino acids, and citrate metabolism (**Figure 5.5; Figure 5.7**). This aligns with previous research showing increased anaerobic metabolism (Hillman 1978a; Katzenback et al. 2014) and mobilization of glucose into the plasma (Malik and Storey 2009). Dehydration tolerance in this animal is thought to be limited by circulatory oxygen delivery such that as dehydration increases, hematocrit is elevated and resting heart rate increases to maintain oxygen delivery (Hillman 1978b, a). At higher levels of dehydration, oxygen delivery is no longer sufficient and anaerobic glycolysis increases as seen by elevated whole body lactate levels (Hillman 1978a).

Similarly, biosynthesis of amino acids in the liver (**Figure 5.7**) correlates with previous observations of increased plasma and liver free amino acid levels in laboratory dehydrated and naturally estivating *X. laevis*, the opposite results seen in starvation alone (Balinsky et al. 1967; Unsworth and Crook 1967). The physiological relevance of increased amino acid levels is not clear, although given that the increase is not uniform across all amino acids it may be a sum of different pathways such as the urea cycle, TCA cycle, and other liver functions involving specific amino acids (Balinsky et al. 1967). This corroboration shows that gene set analysis in the liver closely predicts the physiological and metabolic adaptations of dehydration tolerance that have thus far been determined in *X. laevis*.

5.5.3 Structural proteins only commonality between liver and muscle

Gene set analysis also revealed pathways and processes that had not been previously implicated in *Xenopus* dehydration. Several actin and cytoskeletal related processes and molecular functions were identified as being overrepresented for proteins with corresponding significantly differentially abundant phosphopeptides (**Figure 5.5; Supplementary Figure 5.3; Supplementary Figure 5.4**). While not previously identified as a dehydration tolerance adaptation in *X. laevis*, cytoskeletal reorganization is seen in other dehydration tolerant species ranging from invertebrates to plants (Li et al. 2009; Clark et al. 2009; Cruz DE Carvalho et al. 2014). The loss of body water likely results in reduction in cell volume throughout the body, and thus cell morphology would be impacted which could result in modification to the cytoskeleton to tolerate structural stresses.

Muscle tissue had few enriched terms, although like the liver they include cytoskeletal elements along with regulation of striated muscle contraction (**Figure 5.6**). While the exact functional outcome of these terms being enriched has not been determined, potentially these are adaptations to the structural stresses encountered by cells losing cell volume given that dehydration of muscle tissue negatively impacts muscle contractile function (Lorenzo et al. 2019). It therefore is consistent with the physiology of muscle dehydration to find that regulation of striated muscle contraction was enriched in our dataset. Whether this is a result of dehydration or a response to dehydration is unknown, however it can be speculated that regulation of this process by reversible phosphorylation could be an energy saving mechanism while muscles are dormant, or perhaps a preparatory mechanism for when rehydration occurs.

5.5.4 Regulators of glycolysis phosphorylated during dehydration

Specific enzymes involved in anaerobic glycolysis show how this process may be regulated in response to dehydration. Pyruvate kinase is an ATP producing, regulated step of glycolysis which shows increased activity in *X. laevis* liver during dehydration, potentially due to dephosphorylation (Dawson et al. 2018), which would support anaerobic glycolysis. Our analysis showed that pyruvate kinase was less phosphorylated during dehydration, however this did not reach statistical significance (**Table 5.1**). LDH in the liver also shows altered kinetic and regulatory properties, potentially due to phosphorylation (Katzenback et al. 2014), although we did not identify LDH in our analysis. In this case, LDH is phosphorylated and has lower affinity for its substrates until physiological levels of urea are present which restores kinetic parameters to control levels. The authors proposed that this is a preparatory mechanism for when rehydration occurs and urea is secreted, LDH may remain phosphorylated and therefore less active, shunting pyruvate into the TCA cycle (Katzenback et al. 2014).

The prominence of central metabolic pathways in our gene set enrichment analysis of the liver led us to examine which enzymes in these pathways were differentially phosphorylated in response to dehydration. Particularly interesting was the identification of two PFKFB enzymes (**Table 5.1**). These bifunctional enzymes catalyze the conversion of Fru-6-P to Fru-2,6-P₂ using a PFK2 domain, or the reverse using an FBP2 domain. Fru-2,6-P₂ is the main allosteric activator of phosphofructokinase 1 (PFK1) which catalyzes the first committed ATP utilizing step of glycolysis (the phosphorylation of fructose 6-phosphate to fructose 1,6-bisphosphate) therefore making PFKFB enzymes central regulators of glycolytic flux. Multiple PFKFB enzymes exist, PFKFB1 being the canonical liver isozyme which has previously been identified as a regulatory feature of

dehydration tolerance in the wood frog (Hawkins et al. 2019). Like *X. laevis*, the wood frog survives severe dehydration, in part by funnelling glycogenolytic end products through anaerobic glycolysis (Churchill and Storey 1994). Wood frogs also survives whole body freezing by converting the glycogenolytic end products to cryoprotective glucose and exporting it from the liver to the rest of the body. Both cases, carbon fluxing through glycolysis for energy or export as a cryoprotectant, appear to be controlled at the PFK1 locus and therefore regulated by PFKFB enzymes, potentially by phosphorylation of a serine homologous to human S33 (Vazquez Illanes and Storey 1993; Hawkins et al. 2019).

Our results show that PFKFB1 is indeed dephosphorylated in the liver of *X. laevis* in response to dehydration at S31 and S32 (homologous to human PFKFB1 S33 and S34) (**Table 5.1, Figure 5.8a**). These residues are present in the PFK2 domain of this enzyme (**Figure 5.8b**) and when phosphorylated result in decreased PFK2 activity while increasing FBPase-2 activity (Kurland and Pilkis 1995). Dephosphorylation of PFKFB1 during dehydration would therefore promote PFK2 activity and drive glycolysis. This result is consistent with previous studies showing increased anaerobic glycolysis and regulation of other glycolytic enzymes promoting glycolysis in *X. laevis* liver during dehydration (Hillman 1978a; Katzenback et al. 2014; Dawson et al. 2018).

5.5.5 Hypoxia-inducible PFKFB3 highly upregulated and phosphorylated

PKFB3 was also identified in our results, although this isozyme is not typically expressed in healthy liver tissue. The expression of PFKFB3 is prominent in a number of cancers (Shi et al. 2017) since the *pfkfb3* gene has multiple promoter HREs (binding site for HIF1 α) and is thus induced under hypoxic conditions. This contributes to the high

glycolytic rate of hypoxic tumor cells and has made PFKFB3 a target for pharmaceutical interventions (Wang et al. 2018; Hu et al. 2019; Mondal et al. 2019). Our results indicate that PFKFB3 is phosphorylated during dehydration at S463 (homologous to human S461) (**Figure 5.8a**) which is the residue that activates the PFK2 activity of this enzyme and therefore drives glycolytic flux (Bando et al. 2005). The mechanism of how phosphorylation of this residue increases PFK2 activity has not been studied, however, unlike in PFKFB1 the regulatory residue of PFKFB3 is in close proximity to the binding pocket of the FBPase-2 domain (**Figure 5.8c**). This could result in changes to the kinetics of this enzymes FBPase-2 activity rather than PFK2 activity, increasing the net conversion of Fru-6-P to Fru-2,6-P₂ by inhibiting the reverse reaction. Phosphorylation of PFKFB3, as seen here, would likely contribute to anaerobic glycolysis during the proposed hypoxic conditions seen under high levels of dehydration in *X. laevis*.

Since the *pfkfb3* gene is known to be induced under hypoxic conditions, we measured the multiple transcript variants of this gene and found that all increased (~4-10 fold) during dehydration (**Figure 5.9a,b**). As expected, this upregulation is specific to *pfkfb3* and not a general PFKFB response as the expression of *pfkfb1* actually decreased during dehydration (**Figure 5.9c**). This is likely a result of the eight HREs within the promoter of *pfkfb3* in *X. laevis*, more than the four found in the human *pfkfb3* promoter (Obach et al. 2004) and many more than the two found in the *X. laevis* *pfkfb1* promoter (**Figure 5.9d**). Together these results suggest a possible explanation for how anaerobic glycolysis is elevated in response to high levels of dehydration in *X. laevis*. High levels of dehydration restricts oxygen transport, this leads to a hypoxic cellular environment (Hillman 1978a), which then stimulates transcription of *pfkfb3* (**Figure 5.9b**), which is

then activated by phosphorylation (**Figure 5.8a**), resulting in increased glycolytic flux as evident by elevated lactate levels (Hillman 1978a).

5.5.6 Conclusions

Here we present a phosphoproteomic analysis of *X. laevis* that shows tissue specific responses to dehydration. Liver tissue exhibited a more defined pattern of phosphorylation and dephosphorylation compared to skeletal muscle that is likely the result of the metabolic function this tissue plays during dehydration exposure. Gene set analysis indicated that cellular functions and pathways related to glucose metabolism were highly enriched for differentially phosphorylated proteins in the liver, which lead to the observation that PFKFB3, the hypoxia inducible PFKFB isozyme, increased expression multiple fold and was highly phosphorylated during dehydration. Phosphorylation of PFKFB3 at S463, as seen here, results in increased PFK2 activity of this enzyme, which could explain how anaerobic glycolytic flux increases in response to high levels of dehydration in *X. laevis*. Together, these results highlight the importance of the liver and expand our understanding of dehydration tolerance in this animal.

5.6 References

- Balinsky JB, Choritz EL, Coe CG, van der Schans GS (1967) Amino acid metabolism and urea synthesis in naturally aestivating *Xenopus laevis*. *Comp Biochem Physiol* 22:59–68 . doi: 10.1016/0010-406x(67)90166-1
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Bando H, Atsumi T, Nishio T, et al (2005) Phosphorylation of the 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase/PFKFB3 family of glycolytic regulators in human cancer. *Clin Cancer Res* 11:5784–92 . doi: 10.1158/1078-0432.CCR-05-0149
- Childers CL, Storey KB (2016) Post-translational Regulation of Hexokinase Function and Protein Stability in the Aestivating Frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Chung DJ, Szyszka B, Brown JCL, et al (2013) Changes in the mitochondrial phosphoproteome during mammalian hibernation. *Physiol Genomics* 45:389–399 . doi: 10.1152/physiolgenomics.00171.2012
- Churchill TA, Storey KB (1994) Metabolic responses to dehydration by liver of the wood frog, *Rana sylvatica*. *Can J Zool* 72:1420–1425 . doi: 10.1139/z94-188
- Clark MS, Thorne MAS, Purać J, et al (2009) Surviving the cold: molecular analyses of insect cryoprotective dehydration in the Arctic springtail *Megaphorura arctica* (Tullberg). *BMC Genomics* 10:328 . doi: 10.1186/1471-2164-10-328
- Cowan KJ, Storey KB (2003) Mitogen-activated protein kinases: new signaling pathways functioning in cellular responses to environmental stress. *J Exp Biol* 206:1107–15 . doi: 10.1242/jeb.00220
- Cruz DE Carvalho R, Bernardes DA Silva A, Soares R, et al (2014) Differential proteomics of dehydration and rehydration in bryophytes: evidence towards a common desiccation tolerance mechanism. *Plant Cell Environ* 37:1499–515 . doi: 10.1111/pce.12266
- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- Hawkins LJ, Luu BE, Storey KB (2018) Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. *Gene Reports* 13:192–198 . doi: 10.1016/j.genrep.2018.10.006
- Hawkins LJ, Wang M, Zhang B, et al (2019) Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposures in a freeze tolerant frog. *Comp Biochem Physiol Part D Genomics Proteomics* 30:1–13 . doi: 10.1016/j.cbd.2019.01.009
- Hillman SS (1978a) The roles of oxygen delivery and electrolyte levels in the

- dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Hillman SS (1978b) Some effects of dehydration on internal distributions of water and solutes in *Xenopus laevis*. *Comp Biochem Physiol Part A Physiol* 61:303–307
- Hillman SS, Sommerfeldt RW (1981) Microsphere studies of amphibian systemic blood flow redistribution during dehydration, hypovolemia, and salt load. *J Exp Zool* 218:305–308
- Hu L, Zeng Z, Xia Q, et al (2019) Metformin attenuates hepatoma cell proliferation by decreasing glycolytic flux through the HIF-1 α /PFKFB3/PFK1 pathway. *Life Sci* 239: . doi: 10.1016/j.lfs.2019.116966
- Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9:90–95 . doi: 10.1109/MCSE.2007.55
- Jones E, Oliphant T, Peterson P (2001) SciPy: Open Source Scientific Tools for Python
- Kanehisa M, Sato Y, Furumichi M, et al (2019) New approach for understanding genome variations in KEGG. *Nucleic Acids Res* 47:D590–D595 . doi: 10.1093/nar/gky962
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Kurland IJ, Pilkis SJ (1995) Covalent control of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: insights into autoregulation of a bifunctional enzyme. *Protein Sci* 4:1023–37 . doi: 10.1002/pro.5560040601
- Li A, Benoit JB, Lopez-Martinez G, et al (2009) Distinct contractile and cytoskeletal protein patterns in the Antarctic midge are elicited by desiccation and rehydration. *Proteomics* 9:2788–98 . doi: 10.1002/pmic.200800850
- Liao Y, Wang J, Jaehnig EJ, et al (2019) WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* 47:W199–W205 . doi: 10.1093/nar/gkz401
- Lorenzo I, Serra-Prat M, Yébenes JC (2019) The Role of Water Homeostasis in Muscle Function and Frailty: A Review. *Nutrients* 11:1857 . doi: 10.3390/nu11081857
- Malik AI, Storey KB (2009) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Mole DR, Blancher C, Copley RR, et al (2009) Genome-wide association of hypoxia-inducible factor (HIF)-1 α and HIF-2 α DNA binding with expression profiling of hypoxia-inducible transcripts. *J Biol Chem* 284:16767–75 . doi: 10.1074/jbc.M901790200
- Mondal S, Roy D, Sarkar Bhattacharya S, et al (2019) Therapeutic targeting of PFKFB3 with a novel glycolytic inhibitor PFK158 promotes lipophagy and chemosensitivity in gynecologic cancers. *Int J Cancer* 144:178–189 . doi: 10.1002/ijc.31868
- Obach M, Navarro-Sabaté A, Caro J, et al (2004) 6-Phosphofructo-2-kinase (pfkfb3) gene promoter contains hypoxia-inducible factor-1 binding sites necessary for transactivation in response to hypoxia. *J Biol Chem* 279:53562–70 . doi:

10.1074/jbc.M406096200

- Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine learning in Python. *J Mach Learn Res* 12:2825–2830
- Pellissier F, Glogowski CM, Heinemann SF, et al (2006) Lab assembly of a low-cost, robust SYBR green buffer system for quantitative real-time polymerase chain reaction. *Anal Biochem* 350:310–2 . doi: 10.1016/j.ab.2005.12.002
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45 . doi: 10.1093/nar/29.9.e45
- Ros S, Schulze A (2013) Balancing glycolytic flux: the role of 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatases in cancer metabolism. *Cancer Metab* 1:1 . doi: 10.1186/2049-3002-1-8
- Shi L, Pan H, Liu Z, et al (2017) Roles of PFKFB3 in cancer. *Signal Transduct Target Ther* 2:17044 . doi: 10.1038/sigtrans.2017.44
- Storey KB (2002) Life in the slow lane: Molecular mechanisms of estivation. *Comp Biochem Physiol - A Mol Integr Physiol* 133:733–754 . doi: 10.1016/S1095-6433(02)00206-4
- Storey KB, Storey JM (1990) Metabolic rate depression and biochemical adaptation in anaerobiosis, hibernation and estivation. *Q Rev Biol* 65:145–174
- Su G, Morris JH, Demchak B, Bader GD (2014) Biological network exploration with Cytoscape 3. *Curr Protoc Bioinforma* 47:8.13.1-24 . doi: 10.1002/0471250953.bi0813s47
- Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800 . doi: 10.1371/journal.pone.0021800
- The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* 47:D330–D338 . doi: 10.1093/nar/gky1055
- UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515 . doi: 10.1093/nar/gky1049
- Unsworth BR, Crook EM (1967) The effect of water shortage on the nitrogen metabolism of *Xenopus laevis*. *Comp Biochem Physiol* 23:831–45 . doi: 10.1016/0010-406x(67)90345-3
- Vazquez Illanes MD, Storey KB (1993) 6-Phosphofructo-2-kinase and control of cryoprotectant synthesis in freeze tolerant frogs. *Biochim Biophys Acta* 1158:29–32
- Vertommen D, Herinckx G, Hussain N, et al (2017) Changes in the phosphoproteome of brown adipose tissue during hibernation in the ground squirrel, *Ictidomys tridecemlineatus*. *Physiol Genomics* physiolgenomics.00038.2017 . doi: 10.1152/physiolgenomics.00038.2017
- Wang C, Qu J, Yan S, et al (2018) PFK15, a PFKFB3 antagonist, inhibits autophagy and proliferation in rhabdomyosarcoma cells. *Int J Mol Med* 42:359–367 . doi: 10.3892/ijmm.2018.3599

Waskom M (2018) seaborn

Yang J, Yan R, Roy A, et al (2015) The I-TASSER Suite: protein structure and function prediction. Nat Methods 12:7–8 . doi: 10.1038/nmeth.3213

Table 5.1. Glucose metabolism related enzymes corresponding to identified phosphopeptides. Glycogen metabolism, glycolysis, gluconeogenesis, and TCA cycle enzymes with identified phosphopeptides are presented. FDR corrected p-values in bold are statistically significant ($p < 0.05$).

Pathway	Abbreviation	UniProt	Peptide	Residue	Log2 fold change	FDR corrected p-value
Glycogen metabolism	GSK3 β	Q91757	GEPNV p SYICSR	S215	0.67	0.033
			GEPNV S pYICSR	Y216	0.59	0.063
	PGM	A0A1L8GFG5	AIGGIIL p TASHNPGGPNDFGIK	T115	1.20	0.005
			AIGGIILTA p SHNPGGPNDFGIK	S117	1.20	0.009
Glycolysis/Gluconeogenesis	PFKFB1	Q5U4Z6	RG p SSIPQFTNSPTMIIMVGLPAR	S31	-1.74	0.007
			RGS p SIPQFTNSPTMIIMVGLPAR	S32	-1.67	0.013
	PFKFB3	A0A1L8GTR7	RN p SVTPLASPEPTKK	S463	1.32	0.005
	ALDOB	Q5XHC6	GILAADE p SVGTMGSR	S36	-2.17	0.003
	GAPDH	P51469	VINDNFGIVEGLMTTVHAFT A pTQK	T182	-2.34	0.003
	PK	A0A1L8FDJ9	R p SENMAQLMQDLGPAFVQR	S89	-0.44	0.267
TCA cycle	PDHA1	Q66JA7	YGMGT p SVER	S242	-1.19	0.101
			YHGH p SMSDPGVSYR	S303	-0.28	0.485
			YHGHS M pSDPGVSYR	S305	-0.11	0.785
			YHGHSMSDPGV p SYR	S310	1.02	0.186
			YHGHS M pSDPGV p SYR	S305+S310	-0.49	0.242
	CS	Q5U5A8	TA p SFSESRTEDITPAKK	S455	-1.23	0.044

Abbreviations: **GSK3 β** - Glycogen synthase kinase-3 beta; **PGM** - Phosphoglucomutase; **PFKFB1** - 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1; **PFKFB3** - 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3; **ALDOB** - Fructose-biphosphate aldolase (Liver); **GAPDH** - Glyceraldehyde-3-phosphate dehydrogenase; **PK** - Pyruvate kinase; **PDHA1** - Pyruvate dehydrogenase E1 component subunit alpha; **CS** - ATP-citrate synthase

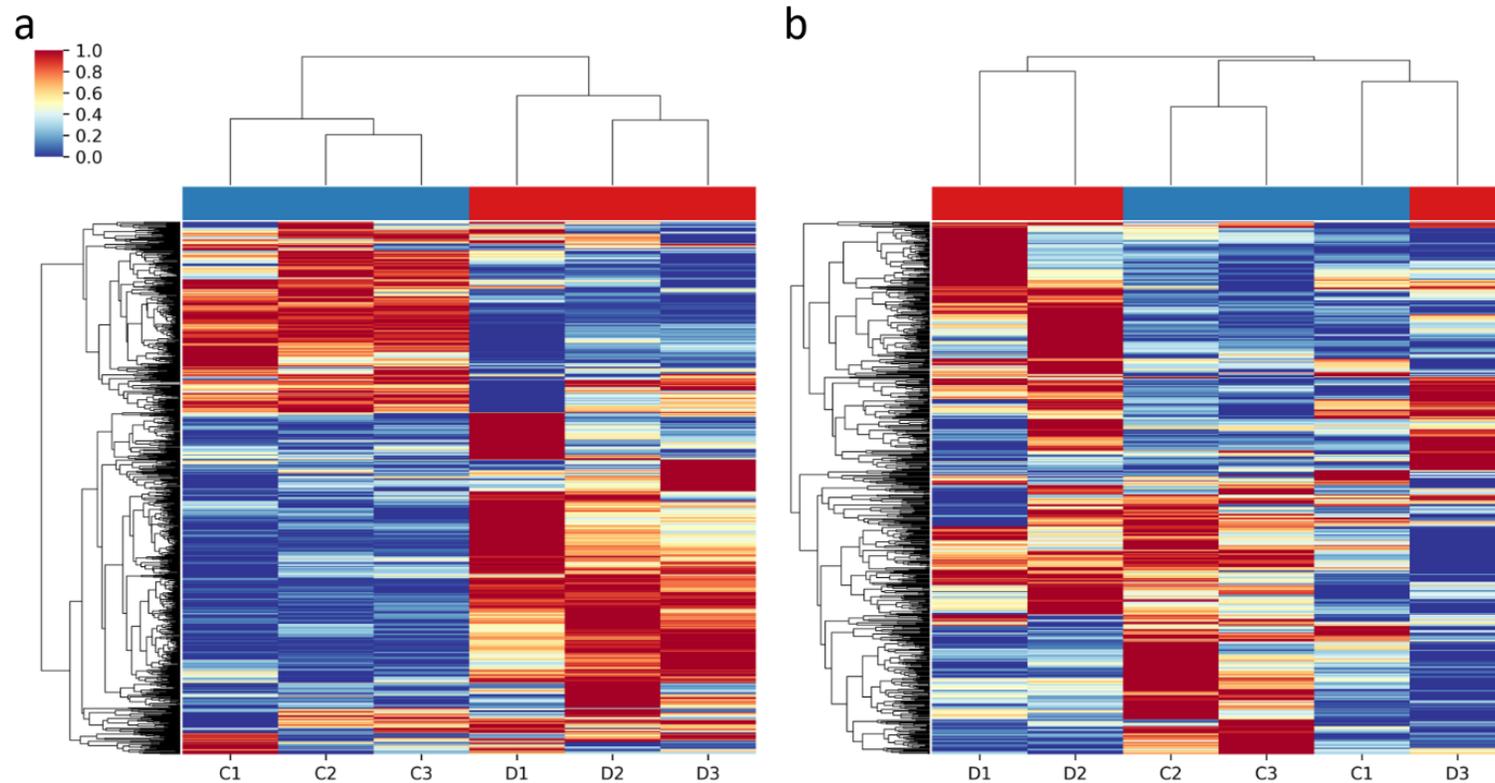


Figure 5.1. Hierarchical clustering analysis of phosphopeptides in the liver and muscle of *Xenopus laevis* in response to dehydration. Clustering was performed on samples from control and dehydrated animals in **a)** liver and **b)** skeletal muscle. Samples (arranged horizontally) were clustered by Euclidean distance between relative quantities of phosphopeptides, and phosphopeptides (arranged vertically) were clustered based on Euclidean distance between each sample. Colors indicate Z-score (standardized quantities) where red is a high Z-score and blue is a low Z-score. The experimental group the sample belongs to is indicated by blue (control) or red (dehydrated) bars above the heatmap.

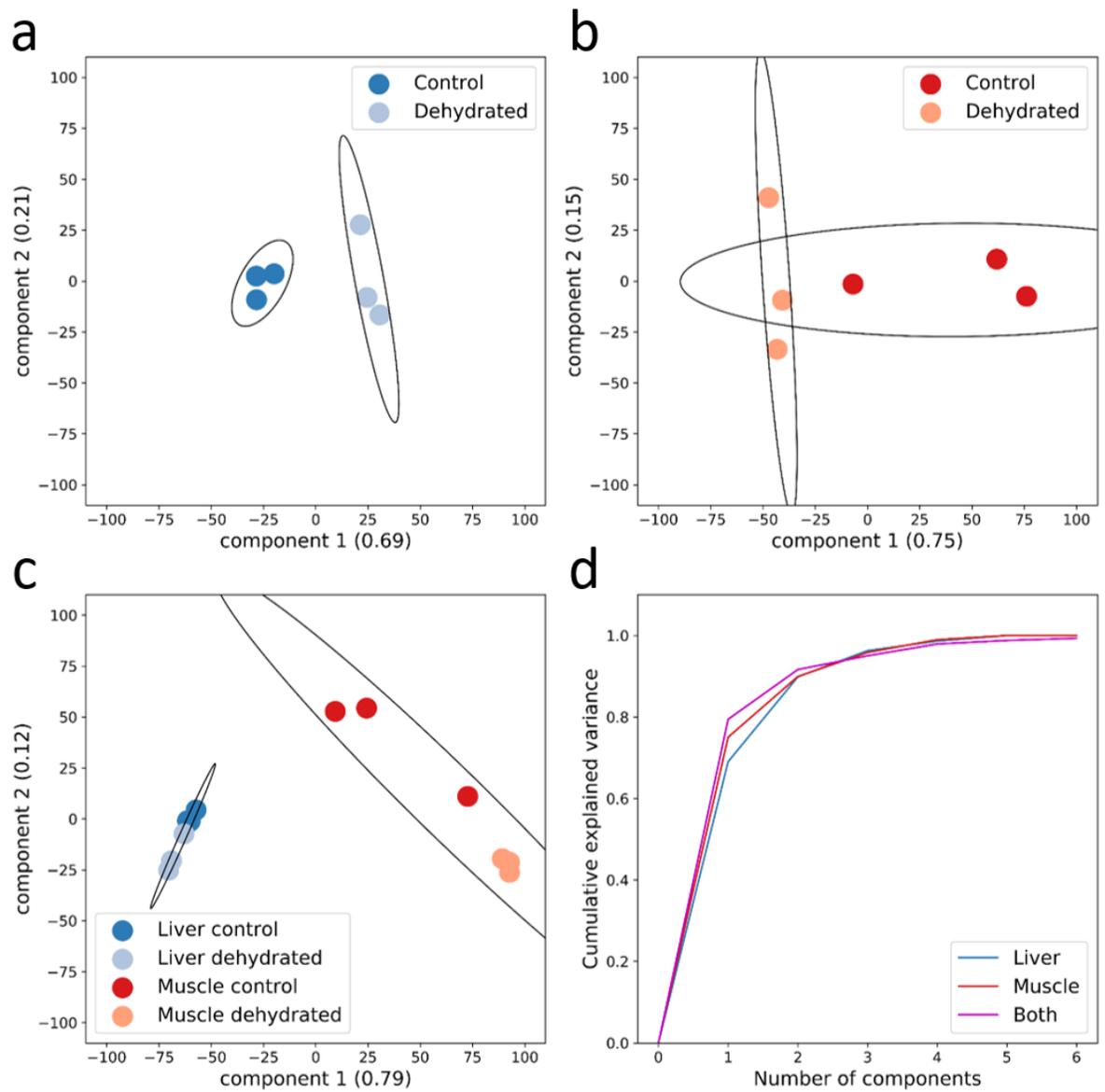


Figure 5.2. Principal component analysis of samples in the liver and muscle of *Xenopus laevis* in response to dehydration. Principle component analysis was performed on phosphopeptide relative quantities in **a**) liver, **b**) skeletal muscle, and **c**) both tissues. The cumulative explained variance of principal components is plotted in **d**).

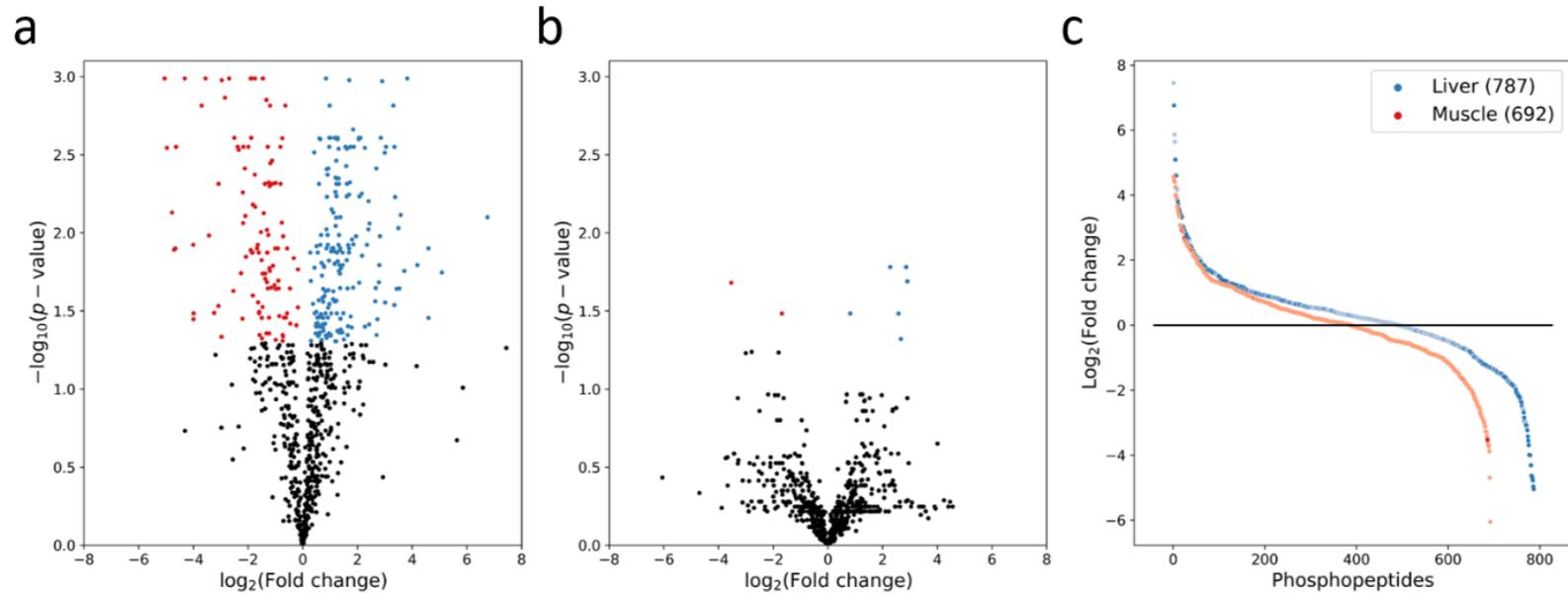


Figure 5.3. Volcano plot of phosphopeptides from liver and muscle of *Xenopus laevis* in response to dehydration. Phosphopeptides that had significantly increased (FDR-corrected p -value < 0.05) and decreased relative abundance in dehydration samples are in blue and red respectively for **a)** liver and **b)** skeletal muscle. Phosphopeptides from liver and skeletal muscle ordered by Log_2 fold-change is shown in **c)**.

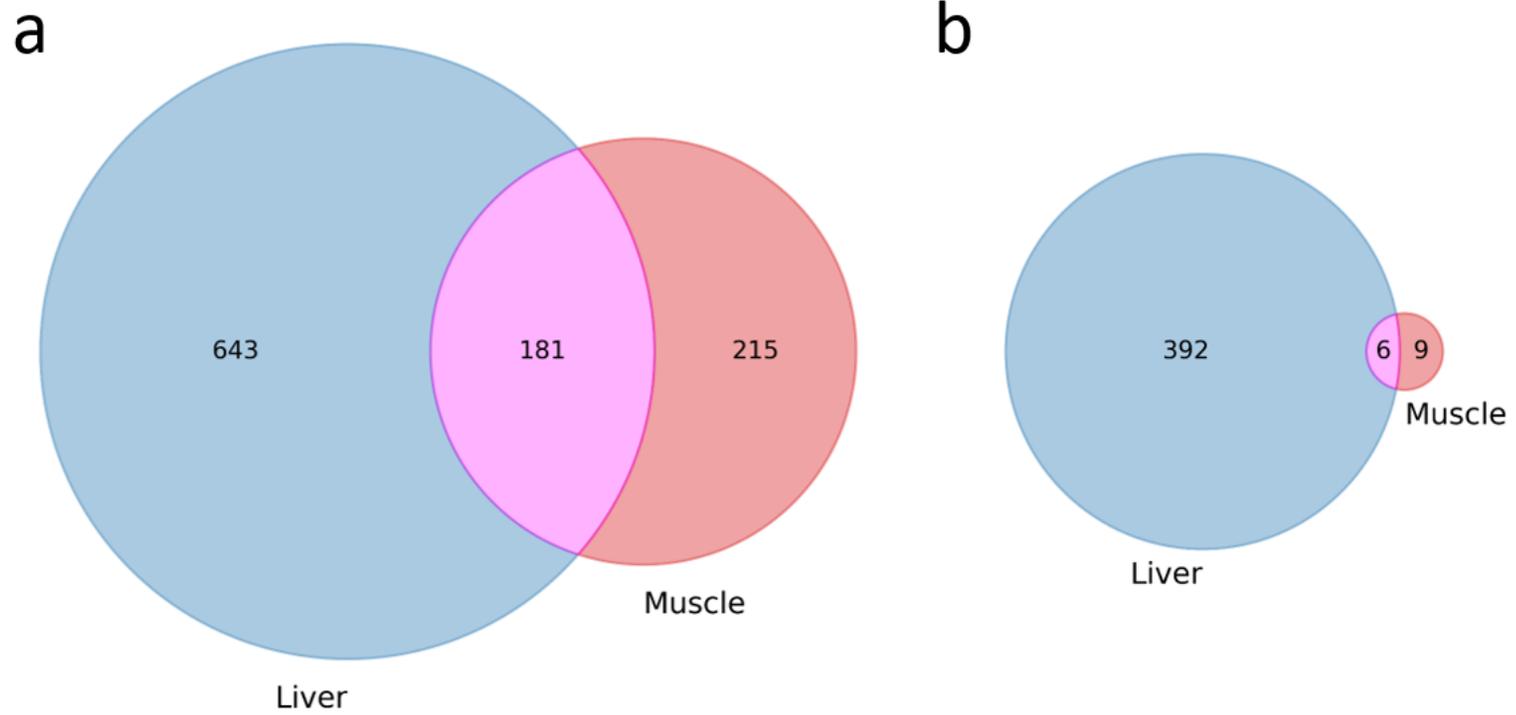


Figure 5.4. Venn diagram of phosphoproteins from liver and muscle of *Xenopus laevis* in response to dehydration. Phosphoproteins with at least one corresponding phosphopeptide identified in liver and/or muscle is shown in **a**), whereas only those with at least one significantly differentially abundant phosphopeptides are shown in **b**).

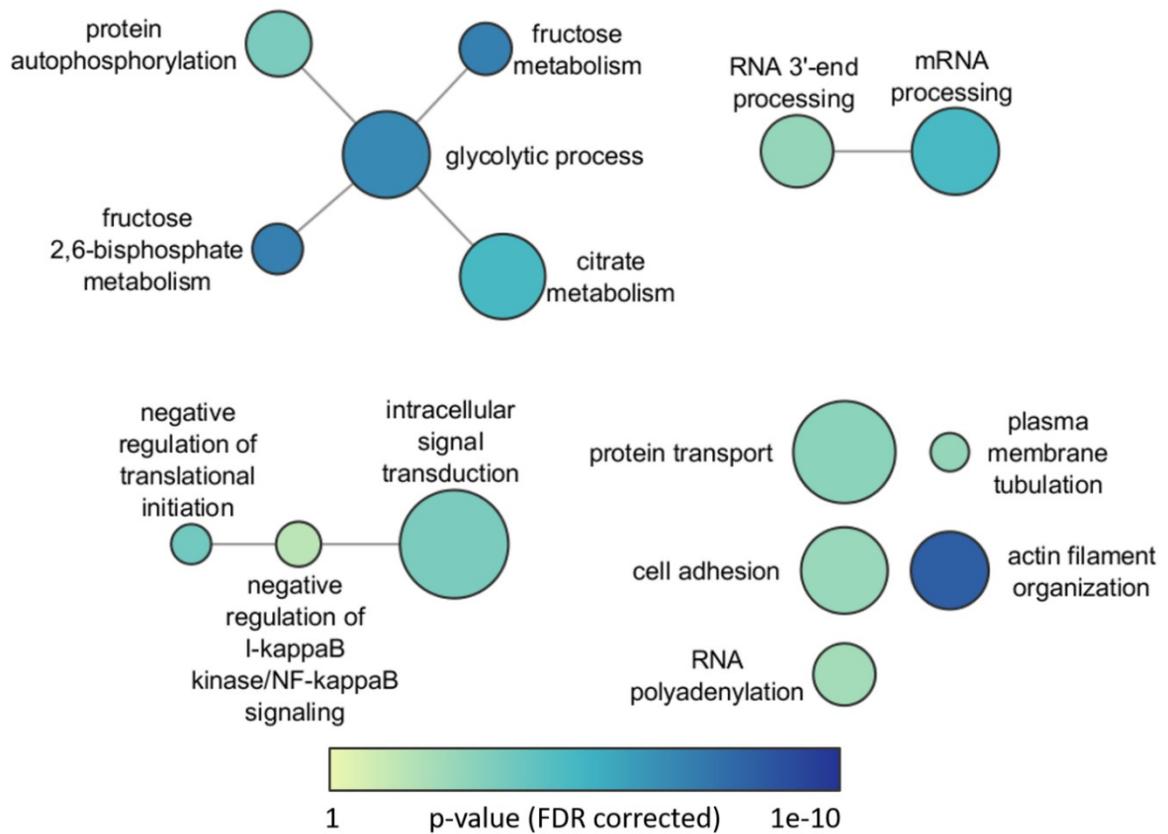


Figure 5.5. Enriched Gene Ontology biological processes semantic relation network in the liver of *Xenopus laevis* exposed to dehydration. Enriched GO biological processes terms were summarized by REVIGO (Supek et al. 2011) and resulting semantic relation network was plotted using Cytoscape. Node size is proportional to the number of genes each term encompasses and the darker the blue the lower the FDR-adjusted p -value. Edges indicate semantic relationship.

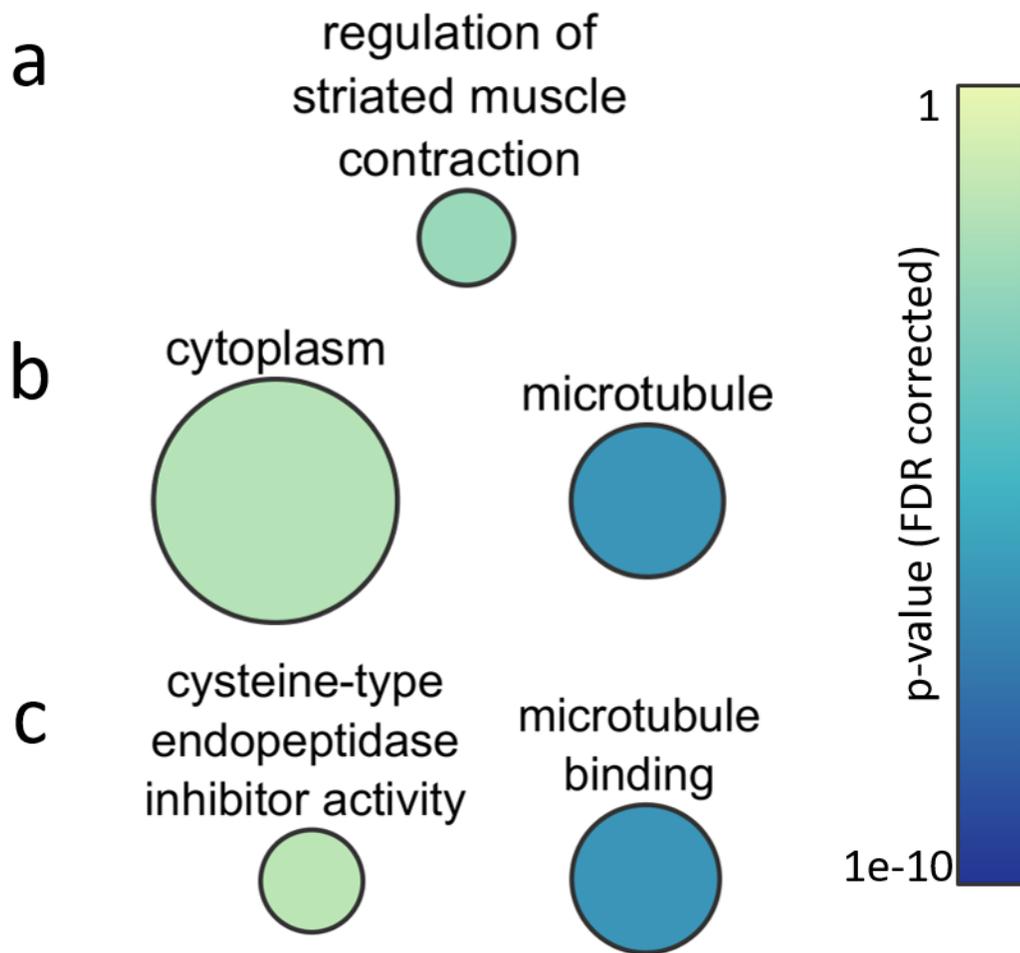


Figure 5.6. Enriched Gene Ontology term semantic relation networks in skeletal muscle of *Xenopus laevis* exposed to dehydration. GO biological processes are in **a)**, cellular compartments in **b)**, and molecular functions in **c)**. All other details as in **Figure 5.5**.

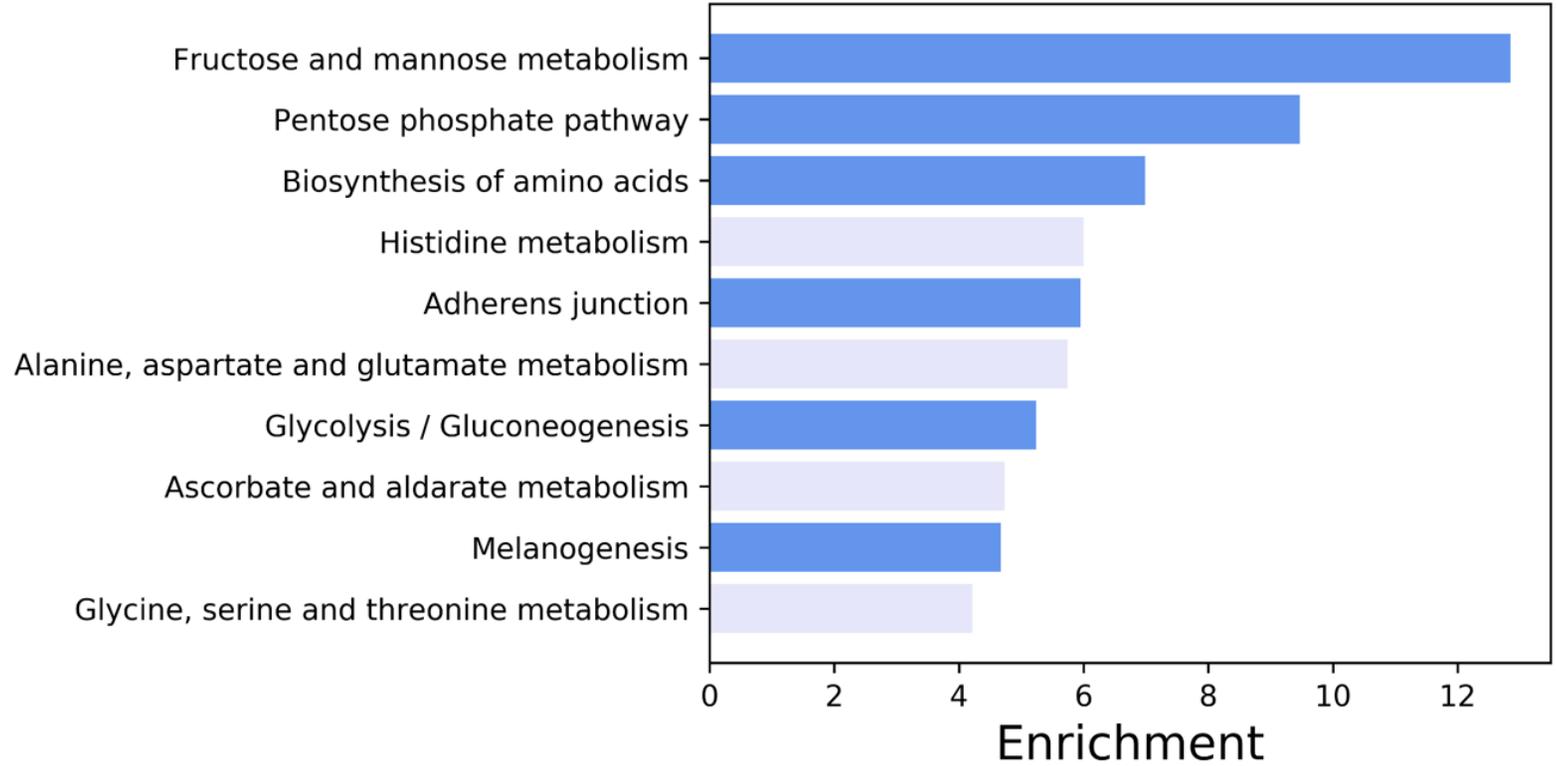


Figure 5.7. Top ten enriched Kyoto Encyclopedia of Genes and Genomes pathways in the liver of *Xenopus laevis* exposed to dehydration. Terms in dark blue are statistically enriched (FDR-corrected p -value < 0.05), terms in light blue are not significantly enriched.

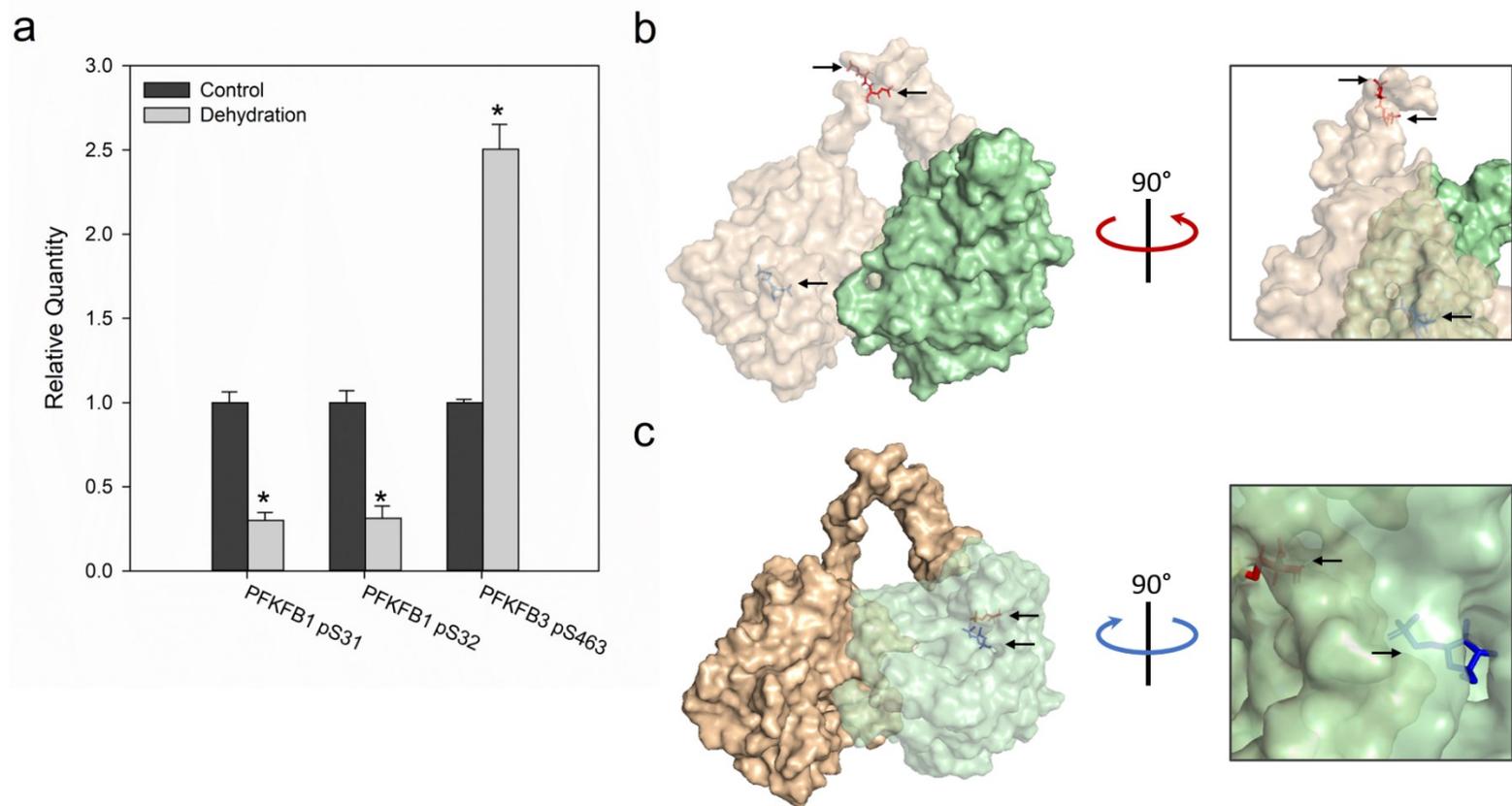


Figure 5.8. Phosphorylation of liver PFKFB enzymes in response to dehydration. PFKFB1 S31, PFKFB1 S32, and PFKFB3 S463 were identified as differentially phosphorylated in the liver during dehydration and are shown in **a**). S31 and S32 are present in the PFK-2 domain (translucent orange) of PFKFB1 as shown in **b**). S463 is present in the FBP-2 domain (translucent green) of PFKFB3 as shown in **c**). Arrows in **b**) and **c**) show phosphorylated serines (red) and substrates (blue).

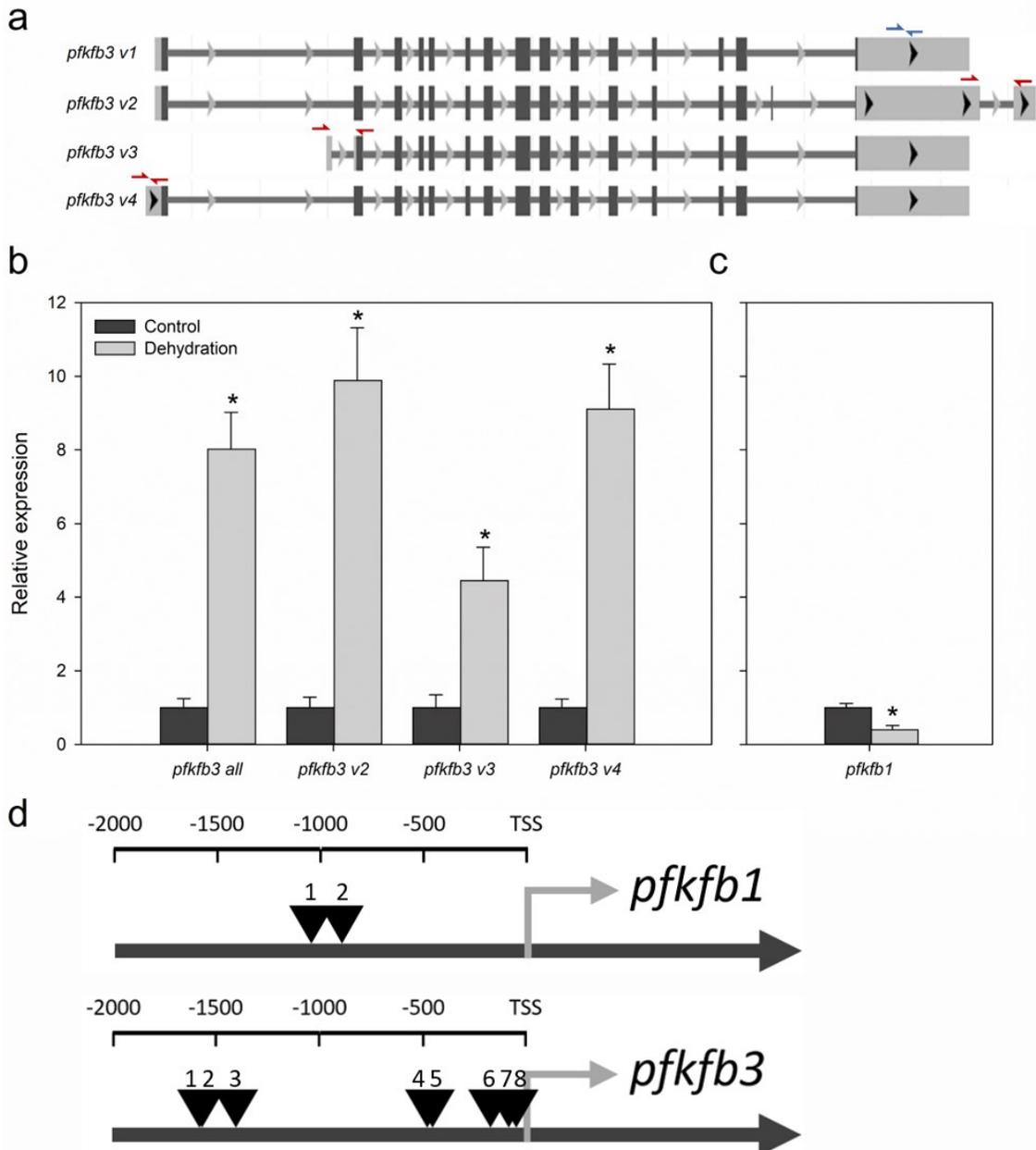
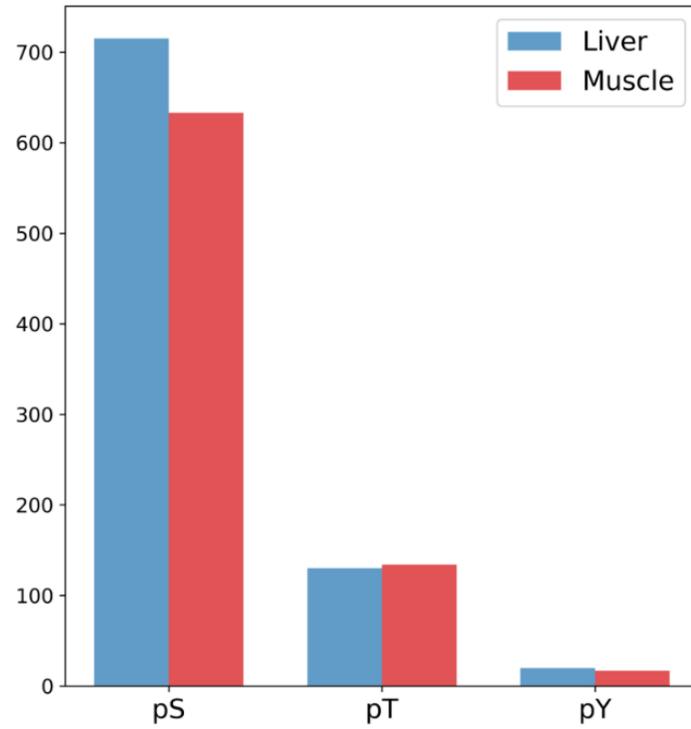
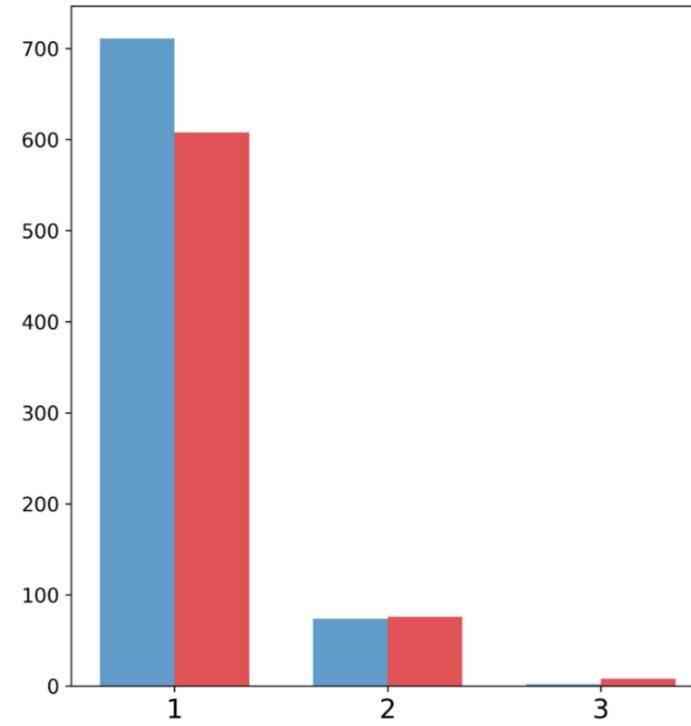


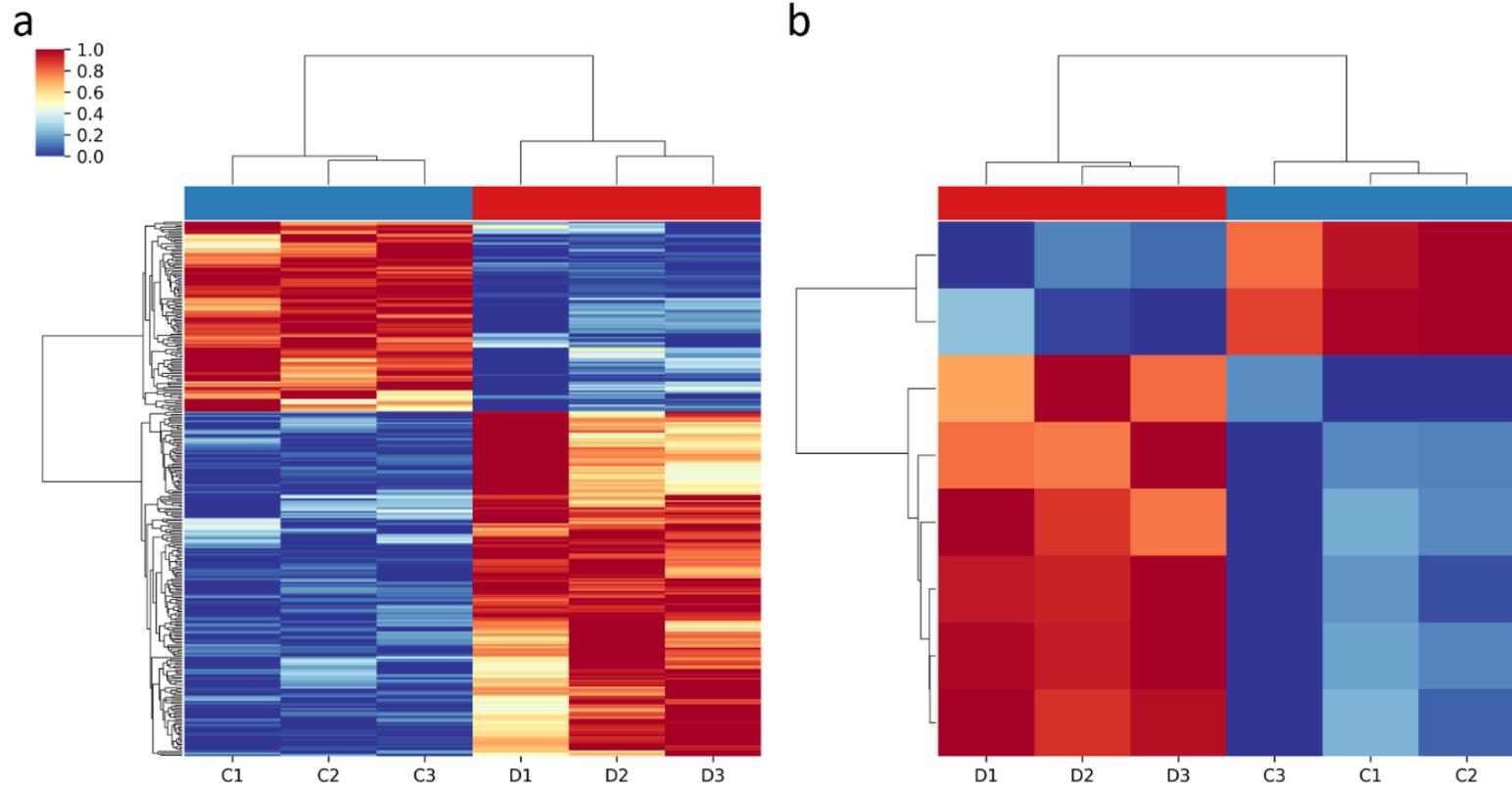
Figure 5.9. Expression of *pfkfb3* and *pfkfb1* in the liver during dehydration. Primers were designed to measure the expression of *pfkfb3* transcript variants as shown by colored arrows in **a**). Blue arrows show the position of forward and reverse primers for measuring all four *pfkfb3* transcript variants, whereas red arrows show primer pairs for measuring specific variants. Primer sequences are in **Supplementary Figure 5.1**. **b**) shows relative expression of *pfkfb3* transcript variants in the liver of control and dehydrated *X. laevis*, whereas relative expression of *pfkfb1* is in **c**). **d**) shows the position of hypoxia responsive elements (HREs, binding sites for HIF1 α) within 2 kbp upstream of the transcription start sites (TSS) of *pfkfb1* and *pfkfb3*.

Supplementary Table 5.1 Primer sequences for qPCR analysis.

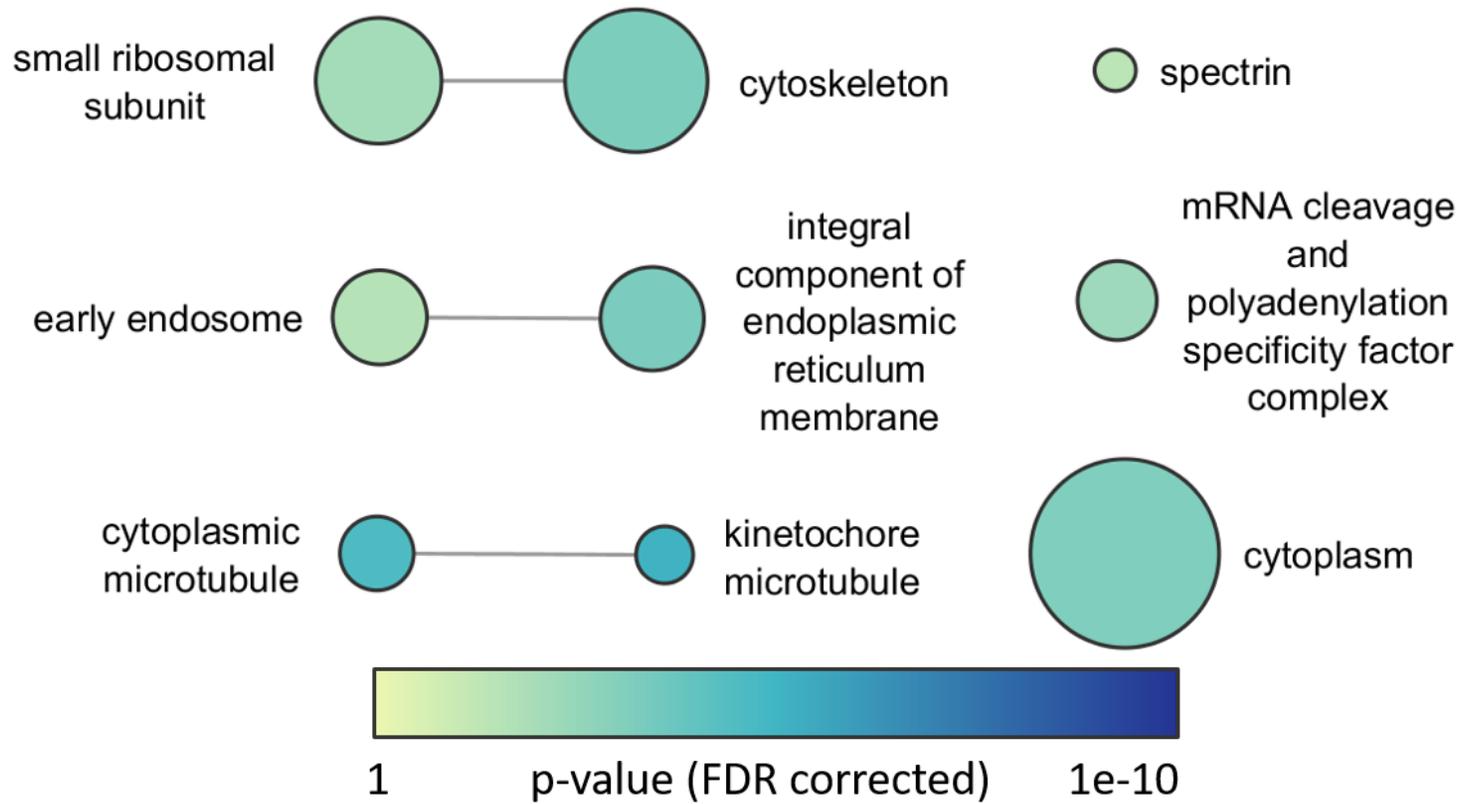
Gene symbol	Accession/GeneID	Forward Primer	Reverse Primer	Product Size
<i>pfkfb3.S all</i>	431861	GTAGAGGGTTTCCCCACCAT	CTTTGCTGATCTCACAAGGGAAA	117
<i>pfkfb3.S v2</i>	XM_018255708	TGGATGCATTACTTCATGTTGCT	CCTGTGCTTCTCCCAGTCAA	115
<i>pfkfb3.S v3</i>	XM_018255709	CCAGCAGAACTTGAAGTGTGC	ACGGGTCAGCTTTTTGGAGAT	120
<i>pfkfb3.S v4</i>	XM_018255710	AGTAATCGAGCGACGGCTTT	GCGAACTAGCAGTACTCCCC	106
<i>pfkfb1.S</i>	NM_001095066	ATGCAGAATGAGGACAGGCTT	CGAGTCTGTTTGAGCTCTCG	119
<i>sdha.L</i>	NM_001090004	TCACGGCTTCTTAGCAGAGC	TGCTGAAGTGTCCCTCCGTC	120
<i>atp5flb.L</i>	XM_018247436	TGCCCGTTTGTCTTGGGTAA	AACAGCCACACAAGTTGGTC	83

a**b**

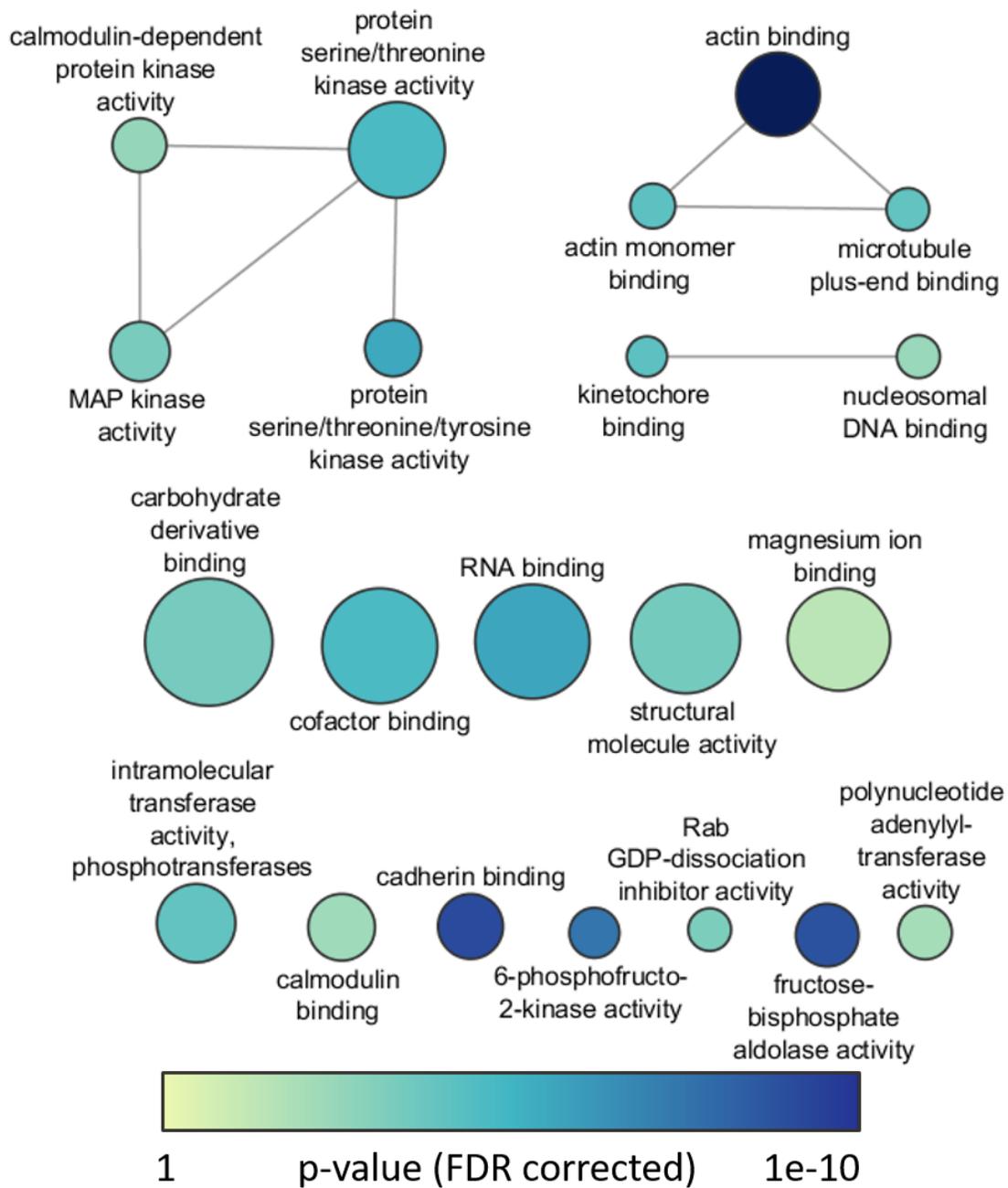
Supplementary Figure 5.1 Summary of phosphopeptide from liver and muscle. **a)** Number of phosphopeptides containing phosphoserine (pS), phosphothreonine (pT), and phosphotyrosine (pY) residues. **b)** number of phosphopeptides containing one, two, or three phosphorylated residues.



Supplementary Figure 5.2 Hierarchical clustering of significantly differentially abundant phosphopeptides from liver and muscle of *Xenopus laevis* exposed to dehydration. Clustering of significantly differential abundance phosphopeptides from **a)** liver and **b)** muscle.



Supplementary Figure 5.3 Semantic relation network of enriched GO cellular compartment terms in the liver of dehydrated *Xenopus laevis*. All other information as in **Figure 5.5**.



Supplementary Figure 5.4 Semantic relation network of enriched GO molecular function terms in the liver of dehydrated *Xenopus laevis*. All other information as in **Figure 5.5**.

Chapter 6: General Discussion

6.1 Discussion

Studying the biological consequences of dehydration is important for a variety of reasons. Athletes, sports institutions, and militaries approach it from a performance perspective, doctors and global health experts from a medical perspective, and agriculturalists from a yield and economic perspective, just to name a few. Organisms that have natural dehydration tolerance therefore are of global interest. These organisms represent nature's solution to the problem of dehydration that have developed over evolutionary time. They therefore serve as resources to kickstart our understanding of dehydration and how to avoid its negative consequences.

For those interested in animal dehydration tolerance, *X. laevis* is the natural choice. This animal survives severe dehydration far beyond other vertebrates, it the most closely related dehydration tolerant animal to humans and is fully characterized as a traditional model organism (**Figure 1.2**). The physiological and metabolic adaptations this animal uses to combat dehydration have been mostly studied in the contexts of the cardiovascular response (Hillman 1978; Hillman and Sommerfeldt 1981) and changes in nitrogen metabolism (Balinsky et al. 1961, 1967). More recently, various studies have taken targeted approaches to studying the biochemical and molecular adaptations of *X. laevis*. For example, since anaerobic glycolysis appears important during dehydration, several studies have chosen specific glycolytic enzymes and examined their kinetics and post-translational modifications (Katzenback et al. 2014; Childers and Storey 2016, 2019; Dawson et al. 2018). Likewise, others have examined the transcripts, proteins, or post-translational modifications of transcription factors (Malik and Storey 2011; Zhang et al. 2018), signaling pathways (Malik and Storey 2009a), antioxidants (Malik and Storey 2009b, 2011), protein chaperones (Luu et al. 2018), and microRNAs (Wu et al. 2013;

Luu and Storey 2015), each revealing potential mechanisms contributing to the dehydration tolerance of this animal.

While each of these previous studies have taken targeted approaches to studying dehydration tolerance, the aim of this thesis was to better situate *X. laevis* as a model organism of dehydration tolerance and to examine multiple regulatory mechanisms of gene expression and function on a large scale. Firstly, I determined appropriate reference genes in multiple tissues of *X. laevis*, creating a stronger foundation for RT-qPCR analysis (**Chapter 2**). Secondly, I quantified global levels of epigenetic mechanisms to determine if there may be genome wide response by these master controllers of transcription (**Chapter 3**). Thirdly, I examined the post-transcriptional regulators, microRNAs, using next-generation sequencing to build on the small-scale studies previously undertaken in this system (**Chapter 4**). Lastly, I showed tissue-specific responses of the phosphoproteome and revealed previously unknown expression and activation of a hypoxia-inducible glycolytic enzyme (**Chapter 5**).

6.2 Reference gene determination

Internal controls are necessary for RT-qPCR due to the sensitivity of sample handling and multiple sample processing steps. The use of traditional housekeeping genes, such as GAPDH and tubulin, has been shown to be unsuitable in a variety of systems, and indeed this may be the case for studying dehydrating *X. laevis*. During dehydration, this animal transitions from aerobic to anaerobic glycolysis and changes in cell volume due to water loss or increased osmotic pressure may affect cytoskeletal elements, thereby potentially involving many common housekeeping genes. In **Chapter 2**, I addressed this issue by combining the three most popular algorithms for validating reference genes (geNorm,

Normfinder, BestKeeper) and evaluated eight candidate genes (*actb.L*, *atp5b1.L*, *canx.L*, *gapdh.S*, *rpl27.L*, *sdha.L*, *tbp.L*, *tbp.S*) in eight tissues (brain, dorsal skin, heart, kidney, liver, lung, skeletal muscle, and ventral skin) of *X. laevis* during dehydration. I found a general trend of consensus between the three algorithms (**Figure 2.3**), and two reference genes was sufficient in all tissues except for kidney where six was optimal (**Figure 2.4; Table 2.2**): *canx.L* and *rpl27.L* in brain, *rpl27.L* and *sdha.L* in dorsal skin, *actb.L* and *rpl27.L* in heart, *actb.L*, *gapdh.S*, *tbp.L*, *atp5f1b.L*, *tbp.S* and *rpl27.L* in kidney, and *atp5f1b.L* in liver, *actb.L* and *rpl27.L* in lung, *gapdh.S* and *rpl27.L* in skeletal muscle, and *rpl27.L* and *canx.L* in ventral skin.

The determination of stable reference genes in these tissues allows not only for more accurate quantification (**Appendix C: RT-qPCR quantification**) of mRNAs in future studies but was also a prerequisite for analysis in later chapters of this thesis. In **Chapter 3** where I examined the response of epigenetic mechanisms to dehydration, these reference genes were used to quantify transcripts of DNMTs, TETs, and TDG in the liver and skeletal muscle. In **Chapter 4** where I performed microRNA-seq, I used the approach from this chapter to identify microRNA that could be used as reference genes to validate the microRNA-seq results with RT-qPCR. Finally, in **Chapter 5**, my phosphoproteomic analysis led me to predict that the transcription of the hypoxia-inducible form of PFK2, *pfkfb3*, would be activated during dehydration which I confirmed using the reference genes found here.

6.2.1 Evaluation of objective 1

Determine suitable reference genes for RT-qPCR analysis of dehydration tolerance in multiple tissues of *Xenopus laevis*

Through the consensus of multiple validation algorithms, stably expressed reference genes were determined for eight tissues of *X. laevis* during dehydration. These reference genes can be used for future RT-qPCR analysis of dehydration tolerance of *X. laevis* and this process of determination can be repeated in other experimental systems.

6.3 Profile of epigenetic mechanisms in response to dehydration

The ability of *X. laevis* to survive dehydration is predicated on changes to the expression of specific genes that confer tolerance. The mechanisms that integrate signals from the environment and regulate gene expression are therefore important to the dehydration tolerance phenotype. For example, transcription factors have shown to be differentially regulated through changes in protein level, phosphorylation status, and DNA binding activity during dehydration which correlate with regulation of their downstream targets (Malik and Storey 2011; Zhang et al. 2018). In **Chapter 3 I** examined a higher level of regulation, epigenetic mechanisms, namely DNA methylation and histone modifications. These mechanisms control genome wide expression and therefore may contribute to dehydration tolerance which requires tissue-specific transcriptional programs.

In this chapter I measured global levels of epigenetic marks, as well as the transcript and activity levels of the enzymes involved in their addition and removal. I showed that although there were no changes in global 5mC levels in skeletal muscle and liver, 5hmC levels increased in the liver (**Figure 3.1**) which may be indicative of a rearrangement of DNA methylation patterns during dehydration. This is consistent with the fact that no change was seen in DNMT activity levels in this tissue (**Figure 3.2**), since a change in net activity of these enzymes is not necessarily a predicate of altering the DNA

methylation patterns given that the patterns are dependant on DNMT-effector protein interactions (Hervouet et al. 2009; Blattler and Farnham 2013). If it is the case that there is a reprogramming of liver DNA methylation patterns, this would suggest changes to transcriptional programming as a result. These speculations of course would need to be validated by further experimentation examining methylation patterns at the gene level and connecting them to transcription genome wide.

The potential liver-specificity of reprogramming is mirrored in **Chapter 5**, as will be discussed, where changes in protein phosphorylation patterns occurred to a much greater extent in the liver than skeletal muscle (**Figure 5.1**, **Figure 5.3**). Our current understanding of the physiological and metabolic consequences of dehydration align with the greater liver-specific responses in these regulatory mechanisms. This organ is metabolically active during dehydration and undergoes multiple metabolic reorganizations. For example, the switch from ammonotelism to ureotelism requires switching nitrogen catabolism pathways. The energy resources necessary for this and other cellular and metabolic functions relies on a redirection from aerobic respiration to anaerobic glycolysis as oxygen transport is hampered. A reprogramming of regulatory mechanisms such as DNA methylation and protein phosphorylation is therefore consistent and expected with the metabolic changes of the liver during dehydration. In contrast, while skeletal muscle has been studied to a lesser extent with respect to dehydration, it likely exists in a state of dormancy or sustainment as we have observed little to no movement by the animals once a water conservation posture is reached.

The upregulation of *tet2* genes in skeletal muscle (**Figure 3.4b**) is interesting since such a large induction (>1400%) did not result in alterations to 5mC or 5hmC levels

(**Figure 3.1**). This result is perhaps a product of the hypoxia-inducibility of these genes (Wu and Zhang 2017), although this does not explain why no increase in transcript levels was seen in the liver (**Figure 3.4a**). Alternatively, the expression of *tet2* genes may be preparative for rehydration, perhaps to act in a myogenic role (Zhong et al. 2017) if disuse atrophy occurs during prolonged dehydration. This preparation and not implementation would be consistent with the phosphoproteomic results in **Chapter 5**, where a high-level analysis revealed that the only cellular process predicted to be affected was *Regulation of Striated Muscle Contraction* (**Figure 5.6a**).

Likewise, of all the histone modifications profiled, H3K4me in the liver was the only modification to show a response to dehydration (**Figure 3.5**). These results also align with the fact that only the liver showed changes in global levels of epigenetic marks as well as the more extensive phosphoproteomic response observed during dehydration.

6.3.1 Evaluation of objective 2

Quantify global changes of DNA methylation and histone modifications in *Xenopus laevis* during dehydration

By measuring several epigenetic markers including DNA methylation and histone modifications I found tissue-specific differences that may reflect epigenetic contributions to dehydration tolerance. These results can be used as a compass for future studies examining gene-specific epigenetic modifications and the role TET enzymes may play in skeletal muscle during dehydration in *X. laevis*.

6.4 Differential expression of microRNA during dehydration

MicroRNAs are cheap and fast-acting post-transcriptional regulators with broad mRNA targeting capabilities, giving the cell a powerful tool for reacting to

environmental stimuli. These short RNA molecules are known to be involved in nearly all cellular functions and are associated with numerous pathologies (Fabris et al. 2016; Martinez et al. 2017; Jużwik et al. 2019). Previous studies have shown differential regulation of microRNAs in *X. laevis*, for example, microRNAs targeting axon guidance and long-term potentiation pathways are downregulated in the brain during dehydration potentially contributing to a neuroprotective phenotype (Luu and Storey 2015). In **Chapter 4**, I examine microRNAs at a larger scale by quantifying the dehydration response of the cardiac microRNAome.

In this chapter I showed that the majority of differentially expressed microRNAs were downregulated in response to dehydration in the heart (**Figure 4.1**), suggesting a removal of negative regulation of the mRNAs they target. This resulted in most of the downstream analysis predicting positive effects on targeted pathways (i.e. less microRNA regulation leading to permissive translational state of members of the pathways). How this trend of downregulation is brought about is unknown at this time. MicroRNA are expressed using the same RNA Pol II machinery that express mRNAs but have a specialized biogenesis pathway that eventually converts their expressed pri-microRNA to the active mature microRNA in the RISC complex (Ha and Kim 2014). This means that the downregulation of these microRNAs could be regulated by cis- or trans-regulatory elements or by the proteins involved in microRNA biogenesis as has been suggested in another stress tolerant anuran (Hadj-Moussa and Storey 2018). Determining how these microRNAs are downregulated would help tie these results into the larger regulatory picture and elucidate the chain of events producing the dehydration tolerance phenotype.

Analysis of pathways, processes, and molecular functions revealed a significant enrichment of microRNA targets whose function involves binding DNA, RNA, nucleosomes, epigenetic regulators (**Figure 4.3, Figure 4.4**), and the *ribosome* and *spliceosome* pathways (**Figure 4.5**). While this study is one of the first forays into the molecular biology of dehydration tolerance in the heart, it is reassuring to note the convergence between these results and results from previous studies (including **Chapter 3**) indicating the involvement of epigenetic, transcriptional, and post-transcriptional effectors during dehydration and other stresses. The results here warrant an investigation into the epigenetics of cardiac dehydration tolerance much like in **Chapter 3** since microRNAs are modelled to relieve negative regulation of targets that bind chromatin, nucleosomes, histones, and HDACs (**Figure 4.3 Figure 4.4**). It may even be the case, that the heart shows a stronger change in epigenetic markers like the liver compared to skeletal muscle based on the enrichment seen here and this tissues strong physiological response to dehydration.

Adding confirmatory evidence to the utility of this type of analysis is the fact that *Cardiac Muscle Contraction* was identified as one of the most enriched pathways for differentially expressed microRNAs (**Figure 4.5, Figure 4.6**). The enrichment analysis performed here is a comparison of the observed number of mRNAs targeted in a pathway compared to the expected number if sampling was random. This means that the targeting of mRNAs in the *Cardiac Muscle Contraction* pathway by the measured microRNAs is statistically different from the same number of microRNAs selected randomly, indicating a coordinating mechanism producing this pattern of microRNA expression in response to

dehydration. Similarly, the fact that the most cardiac-specific KEGG pathway with relevance to the known physiological response emerges lends credence to this analysis.

With respect to the *Cardiac Muscle Contraction* pathway itself, multiple future directions could be explored based on the genes targeted. For example, the opposite function of RyR2 and SERCA2a and their opposing predicted regulation (more targeting of RyR2 and less of SERCA2a) suggests potential changes in calcium cycling dynamics during dehydration (**Figure 4.6**). Likewise, examination of the cross-bridge cycle by TPM and other myofibril proteins may help us understand how the heart is adapted to prolonged dehydration.

The *Glycolysis Gluconeogenesis* pathway was also highly enriched and nearly all genes were predicted to be targeted by microRNAs measured here (**Figure 4.7**). Particularly striking is the fact that all enzymes that lead from glucose to lactate were predicted to be less regulated by microRNAs, except for PGK, which had a slightly negative model coefficient. This is consistent with several studies indicating the importance of anaerobic glycolysis during dehydration in a variety of tissues (Hillman 1978; Katzenback et al. 2014; Childers and Storey 2016, 2019; Dawson et al. 2018) and with results found in **Chapter 5**. Obvious caveats regarding different tissues and the biomolecule under analysis aside, the emergence of *Glucose Gluconeogenesis* from microRNA analysis in the heart (**Figure 4.5, Figure 4.6**) and phosphoproteomic analysis of the liver (**Figure 4.7**) demonstrate the importance of this pathway to the dehydration tolerance phenotype and call for further analysis. Another interesting point of comparison is the apparent lack of response by skeletal muscle (**Chapter 3, Chapter 5**) when compared to cardiac muscle (**Chapter 4**) given how similar these tissues are. This

difference highlights the important physiological differences noted previously; the heart is active throughout dehydration whereas skeletal muscle appears to be physiologically dormant. Of course, various direct comparisons between these two muscle types is needed to fully elucidate their similarities and differences with regard to dehydration tolerance.

6.4.1 Evaluation of objective 3

Profile changes to the cardiac microRNAome in response to dehydration in *Xenopus laevis*

Using an omics approach to study the microRNA response to dehydration in the heart, pathways relevant to the known heart physiology during dehydration emerged. In the most relevant pathways, *Cardiac Muscle Contraction* and *Glycolysis Gluconeogenesis*, nearly all members are predicted to be heavily targeted by dehydration-responsive microRNAs demonstrating a coordinated response to dehydration at the microRNA level.

6.5 Phosphoproteomics reveals *pfkfb3* phosphorylation and expression

Post-translational modification controls numerous parameters of proteins including sub-cellular localization, kinetics, interacting partners, and more. Phosphorylation is particularly well studied, and a wealth of knowledge exists regarding the phosphorylation and dephosphorylation of specific proteins and in many cases specific residues (Hornbeck et al. 2015). Changes in protein phosphorylation status can occur quickly in response to signals from the environment, controlling entire metabolic pathways, therefore it is unsurprising to find phosphorylation changes prominent in studies of environmental stress adaptation (Storey and Storey 2007; Hawkins and Storey 2020). While targeted studies of specific enzymes and proteins have identified functional

changes due to phosphorylation in *X. laevis*, I used phosphoproteomic analysis to investigate large-scale dehydration-induced changes in phosphorylation patterns in **Chapter 5**.

In this chapter I showed strong tissue-level differences in the response of the phosphoproteome to dehydration. Liver showed much more robust and prominent changes in phosphorylation patterns compared to skeletal muscle in multiple types of analyses (**Figure 5.1, Figure 5.2, Figure 5.3, Figure 5.4**) which were not due to differences in the number of detected phosphopeptides in each tissue (**Figure 5.3c**). As indicated previously, these results are consistent with the apparent epigenetic responses to dehydration where liver showed changes in 5hmC and H3K4me levels and skeletal muscle showed no changes to any measured mark (**Figure 3.1, Figure 3.5, Figure 3.6**). Whether these tissue differences in response to dehydration apply to other levels of biological organization is unknown, however, their agreement further supports the importance of the liver during dehydration.

The biological processes associated with changes in phosphorylation showed a cluster related to glycolysis and glucose metabolism which aligns with the known physiological response of the liver, however *Actin filament organization* (**Figure 5.5**) and *Adherens junction* (**Figure 5.7**), both related to structural aspects of the cell, were also identified. Similarly, the *Adherens junction* pathway also emerged in the microRNA analysis of the heart (**Figure 4.5**) even though these were two different tissues with different biomolecules under examination. Dehydration necessarily results in a loss of water from tissues which combined with the accompanying osmotic changes, likely exerts some force on the structural components of the cell and cell-to-cell contacts. The structural

compromising of the cell due to volume-related changes such as dehydration is abundant in the phythological literature (Beck et al. 2007; Oliver et al. 2010), however, membrane and cytoskeletal components of dehydration tolerance have not been studied in *X. laevis*. The fact that cytoskeletal and cell-to-cell components are present in diverse analysis shown here further demonstrates their potential importance to the dehydration tolerance phenotype.

The identification of PFKFB3 phosphorylation (**Figure 5.8**) and upregulation of the hypoxia inducible *pfkfb3* gene (**Figure 5.9**) is a finding not previously seen as an environmental stress adaptation. Glycolysis, and the phosphofructokinase-2 locus in particular, has been the center of intense study in animals that use altered glucose metabolism as part of their response to extreme conditions. For example, the freeze, anoxia, and dehydration tolerant wood frog, *Rana sylvatica*, catabolizes high levels of liver glycogen to produce cryoprotectant glucose that is exported to the blood and transported across the body. While frozen, the liver is subject to hypoxia and eventually anoxia due to freezing of the blood, however, glycolysis is blocked to prevent the catabolism of this cryoprotective glucose (Storey 1987). This raises an interesting question regarding whether *pfkfb3* is expressed in response to hypoxia, as seen in *X. laevis*, even though it would stimulate glycolysis. If *pfkfb3* is expressed in the wood frog, it could still be deactivated by dephosphorylation therefore leaving a mechanism to prevent glycolytic flux. As noted in a study contrasting glycolytic enzyme phosphorylation patterns between wood frogs exposed to freezing, anoxia, and dehydration (Hawkins et al. 2019), there appeared to be phosphorylation patterns that would allow for glucose catabolism during anoxia and dehydration exposure similar to

what is observed in *X. laevis*. These results therefore leave open the possibility that PFKFB3 is involved and potentially unifies the anti-glycolytic response seen during freezing with the catabolism of glucose during dehydration and possibly anoxia exposures.

While further studies characterizing the role of PFKFB3 in dehydration tolerance are needed and will be explored below, the identification of this isozyme demonstrates the utility of discovery-driven phosphoproteomic analyses. To date the approach taken in studying metabolic enzymes in *X. laevis* dehydration tolerance has been to target specific enzymes for purification and characterization. These techniques, while incredibly insightful, would have a difficult time hinting at the expression of a hypoxia-inducible isozyme without *a priori* knowledge. For example, affinity and size-exclusion chromatography may not differentiate between PFKFB1 and PFKFB3 because they use the same substrates and are predicted to be 54.7 kDa and 54.0 kDa respectively. Even when trying to quantify PFKFB3 with techniques such as Western blotting you would run into difficulties since all current commercially available PFKFB3 antibodies have roughly equal epitope identity (~60-70%) with *X. laevis* PFKFB1-4 meaning you could not confidently attribute a measured signal to any of the individual isozymes. Of course, there are techniques that could happen upon exploitable differences between PFKFB1 and PFKFB3, such as ion-exchange chromatography given their differences in isoelectric point, however, proteomic-based analysis removes many of these ambiguities and blind-spots. While the identification of PFKFB3 phosphorylation is an interesting and unanticipated result, there is no doubt that numerous other equally interesting and equally

unanticipated results await in the data presented here. Further context-aware investigation of this data will result in many more exciting dehydration tolerance research avenues.

6.5.1 Evaluation of objective 4

Analyze the liver and skeletal muscle phosphoproteome of *Xenopus laevis* during dehydration

I performed phosphoproteomic analysis of the liver and skeletal muscle and found tissue-specific responses to dehydration. While few phosphorylation changes were observed in skeletal muscle, extensive changes in phosphorylation in the liver occurred indicated significant changes in phenotype. Emerging from the pathway analysis was a hypoxia inducible isozyme of PFKFB that was highly expressed and phosphorylated which may drive anaerobic glycolysis in the liver during dehydration.

6.6 Future directions

While my thesis has broadened our understanding of *X. laevis* dehydration tolerance, possible future directions emerge from the limitations of these studies and the information found within. Firstly, the determination of appropriate reference genes is suitable only for the experimental conditions in which they are determined for. This means that, in the case of the reference genes determined here, they can not be assumed to be stable in other tissues that were not tested and certainly not in other systems of dehydration tolerance. Likewise, different sample preparations may affect what is and is not a suitable reference gene. For this reason, the mRNA reference genes from **Chapter 2** could not be used as reference genes for microRNA validation in **Chapter 4** because they differed in sample preparation, and therefore the validation of microRNAs as reference genes was performed in that chapter. Of course, other types of RNA may be of interest in

the context of dehydration tolerance and may require specialized sample preparations (e.g. long non-coding RNAs, circular RNAs, piwi-interacting RNAs), therefore, to study these, validation of appropriate reference genes would need to be performed on a case-by-case basis. Other normalization strategies such as using external RNA spike-ins or even whole cell spike-ins (Roberts et al. 2014; Taruttis et al. 2017) may be more suitable for complex situations such as normalizing various RNA types in plasma.

The epigenetic marks in **Chapter 3** were measured in a genome-wide manner, and thus interpretation of the results is limited to bulk changes in epigenetic patterns. While large-scale reprogramming of epigenetic marks have been observed in multiple species adapted to environmental stresses (Hawkins and Storey 2020), these large changes occur to a lesser extent in *X. laevis* during dehydration. Of course, these measurements are a summation of all changes across the genome, and therefore within each measurement presented here may be important results at individual genes that do not affect genome-wide measurements due to its limited resolution.

This limitation can be addressed by measuring the epigenetic profile of individual genes. In the case of DNA methylation, it is possible to measure the methylation status at specific genomic loci in a targeted manner using methylation-sensitive high-resolution melt analysis (MS-HRMA; **Appendix E: Methylation-sensitive high-resolution melt analysis**) or in an omics approach using various bisulfite-seq based techniques (Shafi et al. 2018). For histone modifications, methods involving chromatin-immunoprecipitation (ChIP) allow for targeting specific genomic loci (ChIP-qPCR) or all genomic loci (ChIP-seq) and can even be multiplexed to measure many histone modifications across the genome (van Galen et al. 2016). Using methods such as these, individual genes whose

epigenetic profiles change in response to dehydration could be determined. These differentially modified regions (DMRs) would still be correlative measurements, therefore, to determine whether these DMRs contribute to dehydration tolerance or whether they are a result of dehydration tolerance there would need to be an intervention to connect the addition or removal of epigenetic marks to a measurable affect on dehydration tolerance.

The rapid advancement of CRISPR/Cas9 technologies allows for this. While CRISPR/Cas9 is typically associated with its canonical function of genome editing, endonuclease dead Cas9 variants (dCas9) allow for the conjugation with numerous effector domains or proteins to harness the precise targeting of this system to perform a variety of functions at specific genomic loci. Examples include DNMT and TET conjugated dCas9 systems that allow for reprogramming DNA methylation patterns (McDonald et al. 2016; Morita et al. 2016). Furthermore, inducible and tissue-specific CRISPR/Cas9 variants are available (Cao et al. 2016) even in *X. laevis* (DeLay et al. 2018), therefore it is conceivable to develop a strain of *X. laevis* where DNA methylation patterns could be programmed in specific tissues prior to or during exposure to dehydration. Results from experiments like this would provide strong causative evidence for whether DNA methylation contributes to the dehydration tolerance phenotype.

The microRNA-seq analysis in **Chapter 4** identified pathways that were predicted to be regulated in response to dehydration. Perhaps unsurprisingly, the *Cardiac Muscle Contraction* pathway was identified, including members of that pathway that regulate calcium dynamics (**Figure 4.6**). Since calcium is integral to contractility, alterations in calcium dynamics may be a promising direction of study regarding dehydration tolerance.

It is now possible to visualize and characterize calcium dynamics and other parameters *in vivo* in small animals (Jones et al. 2018). Whether or not this technique could be successfully translated to *X. laevis*, especially under conditions of severe dehydration is unknown, however, this is the direction of study that would expand our understanding of how the heart contributes to dehydration tolerance under such stressful conditions. The direct connection with the identified microRNA could be investigated through genetic manipulation (e.g. CRISPR/Cas9 editing of 3'UTR microRNA binding sites), supplemental expression of specific microRNAs, or using anti-microRNA oligonucleotides (Lima et al. 2018). An example application could be to design a heart-specific inducible let-7c-5p expression system. This microRNA was found to be significantly downregulated in the heart (>400% reduction) and is predicted to target multiple members of the *Cardiac Muscle Contraction* pathway (**Supplementary Table 4.3, Figure 4.6**). Restoring the expression of this microRNA during dehydration and examining the effect on cardiomyocyte parameters such as calcium dynamics would clarify the role of this microRNA in dehydration tolerance.

One of the main findings from my phosphoproteomic analysis was that the hypoxia inducible PFKFB3 was highly upregulated and phosphorylated in response to dehydration in the liver and that this could be mediated by HIF1 α (**Figure 5.8, Figure 5.9**). Numerous future directions exist for this result. Firstly, determining how much PFKFB3 contributes to glycolytic flux relative to the canonical liver PFKFB1 is of particular interest. Multiple PFKFB3 specific inhibitors exist due to the prevalence of this isozyme in cancers (Clem et al. 2013). Treatment of *X. laevis* with a PFKFB3 inhibitor, 3PO or PFK15, and measuring dehydration survival and glycolytic parameters such as

lactate production would determine whether PFKFB3 is actually a significant and necessary activator of glycolysis in this context. In a similar vein, these inhibitors could be used in conjunction with fluxomic methods to get a granular view of the flux of metabolic intermediates in glycolysis (Wolak et al. 2012). There is even precedent for the use of fluxomic analysis of environmental stress studies (Tikunov et al. 2014), however, the technical challenges associated with this approach are non-trivial (Tikunov et al. 2013).

Assuming PFKFB3 is a necessary component of dehydration tolerance, the next question is what the causal mechanisms for its induction are. My analysis of the *pfkfb1.S* and *pfkfb3.S* promoters suggested that HIF1 α may be responsible for *pfkfb3.S* induction during dehydration (**Figure 5.9**) which is a mechanism known to occur in cancers, and may actually be the basis for the anti-hepatoma effects of metformin (Hu et al. 2019). This could be more conclusively determined through the use of HIF1 α activators or inhibitors in *X. laevis* (Perry 2020). By treating *X. laevis* with a HIF1 activator such as tilorone (Ratan et al. 2008) it could be observed whether a dehydration-like response occurs with response to PFKFB3. Likewise, administration of a HIF1 inhibitor such as KC7F2 (Narita et al. 2009) or metformin (Hu et al. 2019) during dehydration exposure may block expression of *pfkfb3* and give evidence for a causal connection.

6.7 Final conclusions

Previous studies have taken targeted approaches to studying the dehydration tolerance of *X. laevis*. This thesis has better positioned *X. laevis* as a model for studying dehydration tolerance and examined multiple regulatory mechanisms that could contribute to the dehydration tolerance phenotype. In **Chapter 2** I determined suitable

reference genes for use in RT-qPCR analysis of mRNA transcripts in multiple tissues and used these results as a basis for analysis in later chapters. In **Chapter 3** I measured multiple important epigenetic marks and found tissue-specific differences during dehydration. Specifically, levels of 5hmC rose in the liver but not skeletal muscle, indicating potential epigenetic reprogramming of this tissue. In **Chapter 4** I performed the first microRNA-seq analysis of *X. laevis* dehydration tolerance focusing on the heart. Differential expression of microRNA led to predictions of affected pathways, namely *Cardiac Muscle Contraction* and *Glycolysis Gluconeogenesis*. These two pathways, and their predicted directions of regulation are consistent with the known physiology of *X. laevis* heart during dehydration and provide directions for future studies. Lastly, in **Chapter 5** I performed phosphoproteomic analysis of the liver and skeletal muscle, showing large disparities in responses to dehydration and the emergence of a hypoxia inducible isozyme of PFKFB. Together these results significantly advance our understanding of the molecular biology of dehydration tolerance and provide multiple clear directions for future studies.

6.8 References

- Balinsky JB, Choritz EL, Coe CG, van der Schans GS (1967) Amino acid metabolism and urea synthesis in naturally aestivating *Xenopus laevis*. *Comp Biochem Physiol* 22:59–68 . doi: 10.1016/0010-406x(67)90166-1
- Balinsky JB, Cragg MM, Baldwin E (1961) The adaptation of amphibian waste nitrogen excretion to dehydration. *Comp Biochem Physiol* 3:236–244
- Beck EH, Fettig S, Knake C, et al (2007) Specific and unspecific responses of plants to cold and drought stress. *J Biosci* 32:501–10 . doi: 10.1007/s12038-007-0049-5
- Blattler A, Farnham PJ (2013) Cross-talk between site-specific transcription factors and DNA methylation states. *J Biol Chem* 288:34287–34294 . doi: 10.1074/jbc.R113.512517
- Cao J, Wu L, Zhang S-M, et al (2016) An easy and efficient inducible CRISPR/Cas9 platform with improved specificity for multiple gene targeting. *Nucleic Acids Res* 44:e149 . doi: 10.1093/nar/gkw660
- Childers CL, Storey KB (2016) Post-translational regulation of hexokinase function and protein stability in the aestivating frog *Xenopus laevis*. *Protein J* 35:61–71 . doi: 10.1007/s10930-016-9647-0
- Childers CL, Storey KB (2019) Purification and characterization of a urea sensitive lactate dehydrogenase from skeletal muscle of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B Biochem Syst Environ Physiol* 189:271–281 . doi: 10.1007/s00360-018-1200-3
- Clem BF, O’Neal J, Tapolsky G, et al (2013) Targeting 6-phosphofructo-2-kinase (PFKFB3) as a therapeutic strategy against cancer. *Mol Cancer Ther* 12:1461–1470 . doi: 10.1158/1535-7163.MCT-13-0097
- Dawson NJ, Biggar Y, Malik AI, Storey KB (2018) Increased transcript levels and kinetic function of pyruvate kinase during severe dehydration in aestivating African clawed frogs, *Xenopus laevis*. *Comp Biochem Physiol B Biochem Mol Biol* 0–1 . doi: 10.1016/j.cbpb.2018.01.003
- DeLay BD, Corkins ME, Hanania HL, et al (2018) Tissue-Specific Gene Inactivation in *Xenopus laevis*: Knockout of *lhx1* in the Kidney with CRISPR/Cas9. *Genetics* 208:673–686 . doi: 10.1534/genetics.117.300468
- Fabris L, Ceder Y, Chinnaiyan AM, et al (2016) The Potential of MicroRNAs as Prostate Cancer Biomarkers. *Eur Urol* 70:312–22 . doi: 10.1016/j.eururo.2015.12.054
- Ha M, Kim VN (2014) Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* 15:509–524
- Hadj-Moussa H, Storey KB (2018) Micromanaging freeze tolerance: the biogenesis and regulation of neuroprotective microRNAs in frozen brains. *Cell Mol Life Sci* 75:1–13
- Hawkins LJ, Storey KB (2020) Advances and applications of environmental stress adaptation research. *Comp Biochem Physiol A Mol Integr Physiol* 240:110623 . doi:

10.1016/j.cbpa.2019.110623

- Hawkins LJ, Wang M, Zhang B, et al (2019) Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposures in a freeze tolerant frog. *Comp Biochem Physiol Part D Genomics Proteomics* 30:1–13 . doi: 10.1016/j.cbd.2019.01.009
- Hervouet E, Vallette FM, Cartron P-F (2009) Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. *Epigenetics* 4:487–99 . doi: 10.4161/epi.4.7.9883
- Hillman SS (1978) The roles of oxygen delivery and electrolyte levels in the dehydrational death of *Xenopus laevis*. *J Comp Physiol B* 128:169–175
- Hillman SS, Sommerfeldt RW (1981) Microsphere studies of amphibian systemic blood flow redistribution during dehydration, hypovolemia, and salt load. *J Exp Zool* 218:305–308
- Hornbeck P V., Zhang B, Murray B, et al (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 43:D512-20 . doi: 10.1093/nar/gku1267
- Hu L, Zeng Z, Xia Q, et al (2019) Metformin attenuates hepatoma cell proliferation by decreasing glycolytic flux through the HIF-1 α /PFKFB3/PFK1 pathway. *Life Sci* 239: . doi: 10.1016/j.lfs.2019.116966
- Jones JS, Small DM, Nishimura N (2018) In Vivo Calcium Imaging of Cardiomyocytes in the Beating Mouse Heart With Multiphoton Microscopy. *Front Physiol* 9:969 . doi: 10.3389/fphys.2018.00969
- Jużwik CA, S Drake S, Zhang Y, et al (2019) microRNA dysregulation in neurodegenerative diseases: A systematic review. *Prog Neurobiol* 182:101664 . doi: 10.1016/j.pneurobio.2019.101664
- Katzenback BA, Dawson NJ, Storey KB (2014) Purification and characterization of a urea sensitive lactate dehydrogenase from the liver of the African clawed frog, *Xenopus laevis*. *J Comp Physiol B* 184:601–11 . doi: 10.1007/s00360-014-0824-1
- Lima JF, Cerqueira L, Figueiredo C, et al (2018) Anti-miRNA oligonucleotides: A comprehensive guide for design. *RNA Biol* 15:338–352 . doi: 10.1080/15476286.2018.1445959
- Luu BE, Storey KB (2015) Dehydration triggers differential microRNA expression in *Xenopus laevis* brain. *Gene* 573:64–9
- Luu BEE, Wijenayake S, Malik AII, Storey KBB (2018) The regulation of heat shock proteins in response to dehydration in *Xenopus laevis*. *Cell Stress Chaperones* 23:45–53 . doi: 10.1007/s12192-017-0822-9
- Malik AI, Storey KB (2011) Transcriptional regulation of antioxidant enzymes by FoxO1 under dehydration stress. *Gene* 485:114–9 . doi: 10.1016/j.gene.2011.06.014
- Malik AI, Storey KB (2009a) Activation of extracellular signal-regulated kinases during dehydration in the African clawed frog, *Xenopus laevis*. *J Exp Biol* 212:2595–603
- Malik AI, Storey KB (2009b) Activation of antioxidant defense during dehydration stress

- in the African clawed frog. *Gene* 442:99–107 . doi: 10.1016/j.gene.2009.04.007
- Martinez EC, Lilyanna S, Wang P, et al (2017) MicroRNA-31 promotes adverse cardiac remodeling and dysfunction in ischemic heart disease. *J Mol Cell Cardiol* 112:27–39 . doi: 10.1016/j.yjmcc.2017.08.013
- McDonald JI, Celik H, Rois LE, et al (2016) Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. *Biol Open* 5:866–874 . doi: 10.1242/bio.019067
- Morita S, Noguchi H, Horii T, et al (2016) Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. *Nat Biotechnol* 34:1060–1065 . doi: 10.1038/nbt.3658
- Narita T, Yin S, Gelin CF, et al (2009) Identification of a novel small molecule HIF-1 α translation inhibitor. *Clin Cancer Res* 15:6128–36 . doi: 10.1158/1078-0432.CCR-08-3180
- Oliver MJ, Cushman JC, Koster KL (2010) Dehydration Tolerance in Plants. In: Sunkar R (ed) *Plant Stress Tolerance*. Humana Press, pp 3–24
- Perry SF (2020) Personal communication
- Ratan RR, Siddiq A, Aminova L, et al (2008) Small molecule activation of adaptive gene expression: tilorone or its analogs are novel potent activators of hypoxia inducible factor-1 that provide prophylaxis against stroke and spinal cord injury. *Ann N Y Acad Sci* 1147:383–94 . doi: 10.1196/annals.1427.033
- Roberts TC, Coenen-Stass AML, Wood MJA (2014) Assessment of RT-qPCR normalization strategies for accurate quantification of extracellular microRNAs in murine Serum. *PLoS One* 9: . doi: 10.1371/journal.pone.0089237
- Shafi A, Mitrea C, Nguyen T, Draghici S (2018) A survey of the approaches for identifying differential methylation using bisulfite sequencing data. *Brief Bioinform* 19:737–753 . doi: 10.1093/bib/bbx013
- Storey KB (1987) Glycolysis and the regulation of cryoprotectant synthesis in liver of the freeze tolerant wood frog. *J Comp Physiol B* 157:373–380 . doi: 10.1007/BF00693364
- Storey KB, Storey JM (2007) Tribute to P. L. Lutz: putting life on 'pause' --molecular regulation of hypometabolism. *J Exp Biol* 210:1700–14 . doi: 10.1242/jeb.02716
- Taruttis F, Feist M, Schwarzfischer P, et al (2017) External calibration with *Drosophila* whole-cell spike-ins delivers absolute mRNA fold changes from human RNA-Seq and qPCR data. *Biotechniques* 62:53–61 . doi: 10.2144/000114514
- Tikunov AP, Stoskopf MK, Macdonald JM (2014) Fluxomics of the Eastern Oyster for Environmental Stress Studies. *Metabolites* 4:53–70 . doi: 10.3390/metabo4010053
- Tikunov AP, Winnike JH, Tech K, et al (2013) Fluxomics by NMR Spectroscopy from Cells to Organisms Focusing on Liver. *Curr Metabolomics* 1:128–159 . doi: 10.2174/2213235X11301020003
- van Galen P, Viny AD, Ram O, et al (2016) A Multiplexed System for Quantitative

- Comparisons of Chromatin Landscapes. *Mol Cell* 61:170–80 . doi: 10.1016/j.molcel.2015.11.003
- Wolak J, Rahimi-Keshari K, Jeffries RE, et al (2012) Noninvasive Fluxomics in Mammals by Nuclear Magnetic Resonance Spectroscopy. In: *Handbook of Metabolomics*. pp 321–392
- Wu CW, Biggar KK, Storey KB (2013) Dehydration mediated microRNA response in the African clawed frog *Xenopus laevis*. *Gene* 529:269–75
- Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 18:517–534 . doi: 10.1038/nrg.2017.33
- Zhang Y, English SG, Storey KB (2018) Regulation of nuclear factor of activated T cells (NFAT) and downstream myogenic proteins during dehydration in the African clawed frog. *Mol Biol Rep* 45:751–761 . doi: 10.1007/s11033-018-4214-8
- Zhong X, Wang Q-Q, Li J-W, et al (2017) Ten-Eleven Translocation-2 (Tet2) Is Involved in Myogenic Differentiation of Skeletal Myoblast Cells in Vitro. *Sci Rep* 7:43539 . doi: 10.1038/srep43539

Appendices

Appendix A: List of publications

1. Thorne MAS, Britovšek NK, **Hawkins LJ**, Lilley KS, Storey KB (2020) Proteomics of intracellular freezing survival. PLOS ONE. <https://doi.org/10.1371/journal.pone.0233048>
2. **Hawkins LJ**, Storey KB (2020) MicroRNA expression in the heart of *Xenopus laevis* facilitates metabolic adaptation to dehydration. Genomics. <https://doi.org/10.1016/j.ygeno.2020.04.003>
3. **Hawkins LJ**, Storey KB (2020) Advances and applications of environmental stress adaptation research. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology. 240, 110623. <https://doi.org/10.1016/j.cbpa.2019.110623>
4. **Hawkins LJ***, Hadj-Moussa H*, Nguyen VC, Pamerter ME, Storey KB (2019) Naked mole rats activate neuroprotective proteins during hypoxia. Journal of Experimental Zoology Part A. 331(10), 571-576. <https://doi.org/10.1002/jez.2321>
5. **Hawkins LJ**, Wang M, Xiao Q, Wang H, Storey KB (2019) Glucose and urea metabolic enzymes are differentially phosphorylated during freezing, anoxia, and dehydration exposure in a freeze tolerant vertebrate. Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics. 30. 1-13. <https://doi.org/10.1016/j.cbd.2019.01.009>
6. Zhang J*, **Hawkins LJ***, Storey KB (2019) DNA methylation and regulation of DNA Methyltransferases in a freeze tolerant vertebrate. Biochemistry and Cell Biology. <https://doi.org/10.1139/bcb-2019-0091>
7. **Hawkins LJ***, Al-attar R*, Storey KB (2018) Transcriptional regulation of metabolism in disease: from transcription factors to epigenetics. PeerJ. 6:e5062. <https://doi.org/10.7717/peerj.5062>
8. **Hawkins LJ**, Luu BE, Storey KB (2018) Selection of reference genes for accurate RT-qPCR analysis of dehydration tolerance in *Xenopus laevis*. Gene Reports. 13. 192-198. <https://doi.org/10.1016/j.genrep.2018.10.006>
9. **Hawkins LJ**, Storey KB (2018) Histone methylation in the freeze tolerant wood frog (*Rana sylvatica*). Journal of Comparative Physiology B. 188(1), 113-125. <https://doi.org/10.1007/s00360-017-1112-7>
10. Wijenayake S, **Hawkins LJ**, Storey KB (2018) Regulation of histone lysine methylation in response to anoxia tolerance in a freshwater turtle, *Trachemys scripta elegans*. Gene. 649(5), 50-57. <https://doi.org/10.1016/j.gene.2018.01.086>
11. Rouble AN, **Hawkins LJ**, Storey KB (2018) Roles for lysine acetyltransferases during Mammalian hibernation. Journal of Thermal Biology. 74, 71-76. <https://doi.org/10.1016/j.jtherbio.2018.03.013>
12. **Hawkins LJ**, Storey KB (2017) Improved high-throughput quantification of luminescent microplate assays using a common Western-blot imaging system. MethodsX. 4, 413-422. <https://doi.org/10.1016/j.mex.2017.10.006>

* Equal Contribution

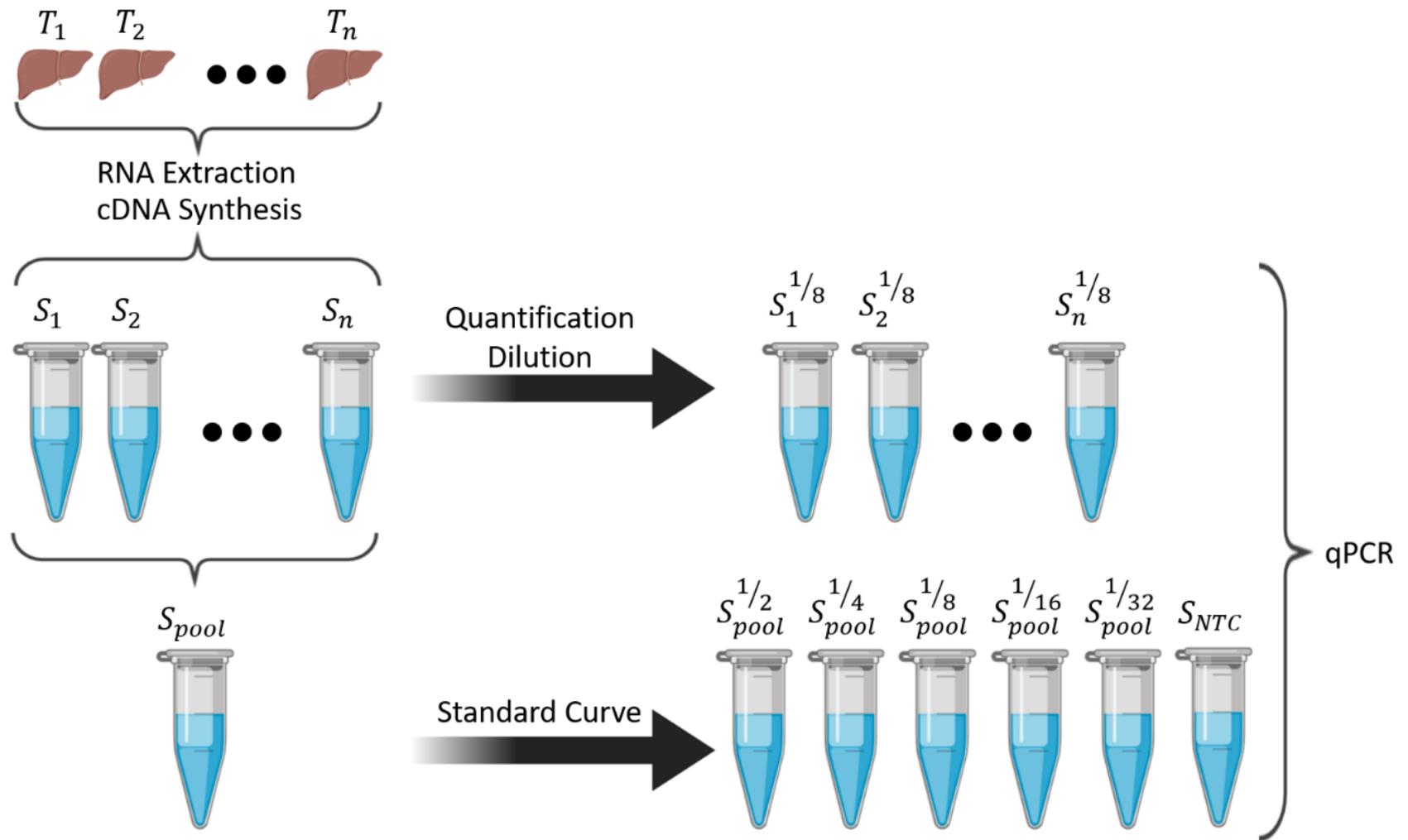
Appendix B: Communications at scientific meetings

1. **Hawkins LJ***, Storey KB (2019). Phosphorylation of glycolysis and urea cycle enzymes in response to multiple stresses in a freeze tolerant vertebrate. The 10th International Congress of Comparative Physiology and Biochemistry. Ottawa, Ontario, Oral presentation
2. **Hawkins LJ**, Storey KB* (2018). Changes in histone methyltransferases during freezing stress in the wood frog. 13th International Conference on Genomics. Shenzhen, China. Poster Presentation
3. **Hawkins LJ***, Storey KB (2018). Extending the applications of common laboratory equipment: A case study. Ottawa-Carleton Chemistry Institute OCCI Day 2018. Ottawa, Ontario. Oral presentation.
4. **Hawkins LJ***, Storey KB (2018). Extending the applications of common laboratory equipment: A case study. 15th Annual Ottawa-Carleton Institute of Biology (OCIB) Symposium. Ottawa, Ontario. Oral presentation.
5. **Hawkins LJ***, Storey KB (2017). Epigenetic regulators in the dehydration tolerant *Xenopus laevis*. Canadian Society for Molecular Biosciences 60th Annual Conference. Ottawa, Ontario. Poster presentation.
6. **Hawkins LJ**, Storey KB* (2017). Changes in histone methyltransferases during freezing stress in the wood frog. Canadian Society of Zoologists, Satellite Symposium: 50 years of comparative biochemistry – the legacy of Peter Hochachka. University of Manitoba, Winnipeg, Manitoba. Poster presentation.
7. Al-attar R*, **Hawkins LJ**, Bedard N, Wing SS, Storey KB (2017). Differential regulation of microRNAs in the liver of USP19 null mice during fasting. Toronto RNA Enthusiasts Day, Toronto, Ontario. Poster presentation.
8. **Hawkins LJ***, Storey KB (2016). Changes in histone methyltransferases during freezing stress in the wood frog. *Cryobiology* 73, 437-438. 53rd Annual Meeting of the Society for Cryobiology. Ottawa, Ontario. Poster presentation.
9. **Hawkins LJ***, Storey KB (2016). Changes in histone methyltransferases during freezing stress in the wood frog. 13th Annual Ottawa-Carleton Institute of Biology (OCIB) Symposium. Ottawa, Ontario. Poster presentation.
10. **Hawkins LJ***, Storey KB (2015). Natural freeze tolerance and histone methyltransferases in the wood frog. 12th Annual Ottawa-Carleton Institute of Biology (OCIB) Symposium. Ottawa, Ontario. Poster presentation.
11. **Hawkins LJ***, Leonard M (2013). Ambient noise, alarm call structure, and recruitment in adult Tree Swallows (*Tachycineta bicolor*). 27th Annual Cameron Conference for Biology and Marine Biology Honors Students. Halifax, Nova Scotia. Poster presentation.

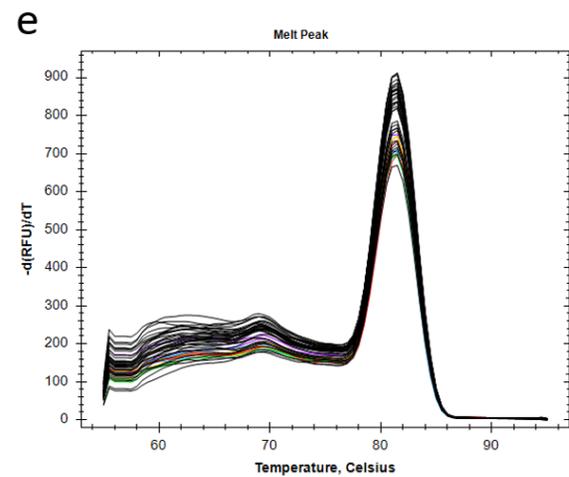
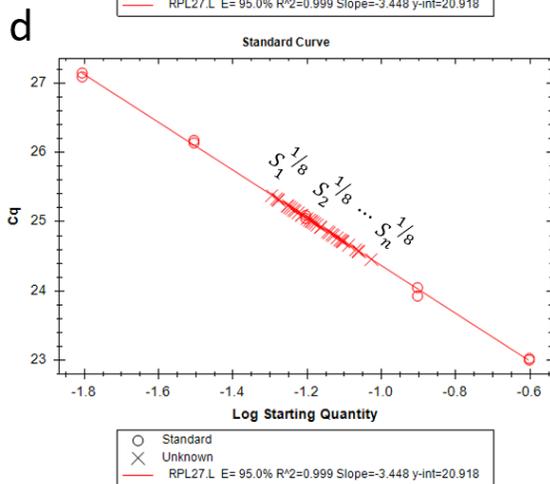
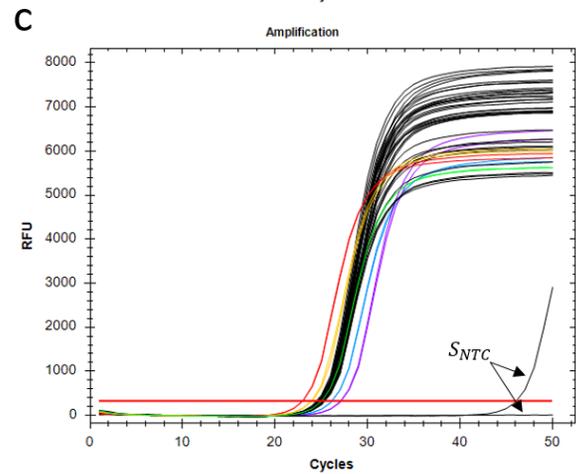
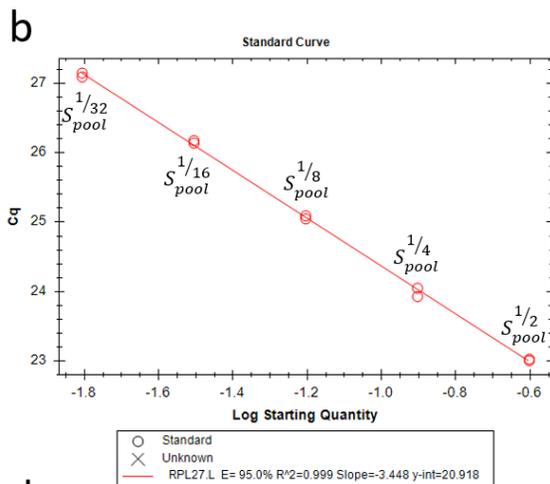
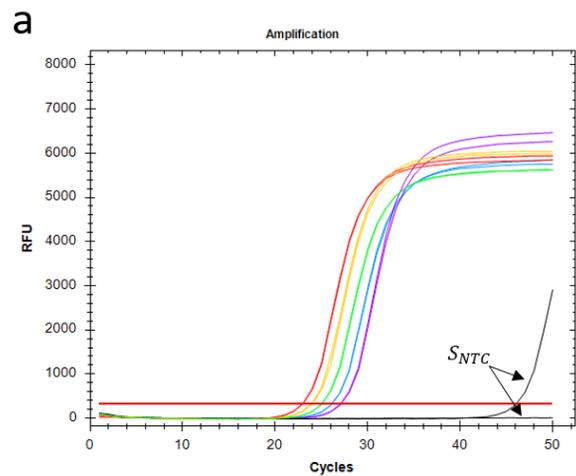
* Presenter

Appendix C: RT-qPCR quantification

All mRNAs and microRNAs quantified using RT-qPCR in this thesis was done in the following way. Tissues ($T_1 - T_n$) from individual animals undergo RNA extraction and cDNA synthesis to produce cDNA samples ($S_1 - S_n$) (**Appendix Figure 1**). A pooled sample (S_{pool}) is produced from all cDNA samples which is used to produce a 2-fold serial dilution standard curve of at least five points ($S_{pool}^{1/2} - S_{pool}^{1/32}$) and a no template control (S_{NTC}). All cDNA samples to be quantified are diluted to the same dilution corresponding to the middle of the standard curve (1/8 dilution in the case presented in **Appendix Figure 1**). After amplification the standard curve samples show even spacing, an efficiency between 90-110%, and late or no amplification of the NTC sample (**Appendix Figure 2a,b**). When the amplification curves of the diluted samples of interest ($S_1 - S_n$) are overlaid, their C_q surround their corresponding dilution sample from the standard curve (1/8 in **Appendix Figure 2c,d**). Melt curve analysis indicates that the same single product is amplified in all samples (**Appendix Figure 2e**).



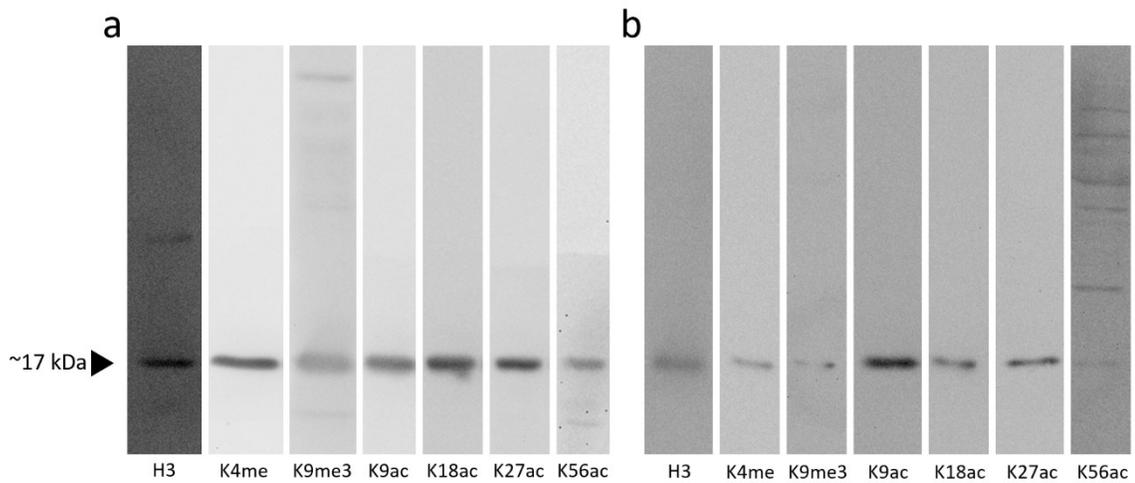
Appendix Figure 1. Sample schema for quantification of mRNAs and microRNAs using RT-qPCR.



Appendix Figure 2. Example RT-qPCR results. a) Amplification of 2-fold serial dilution standard curve (rainbow curves) shows expected equidistant C_q values (red line) and NTC samples amplifying much later than standard curve samples or not at all. b) Plot of log of the dilution factor versus C_q shows reaction efficiency of 95.0% and R^2 of 0.999. c) Overlay of amplification of samples of interest. d) Samples of interest fall near middle of standard curve as expected from dilution. e) Melt curve analysis indicates a single product was amplified in all samples.

Appendix D: Full lane Western blot images

Displayed below are full lane images for the Western blots presented in this thesis. All lanes chosen for **Figure 3.5 (Appendix Figure 3a)** corresponds to the first control sample from the liver, likewise, all lanes chosen for **Figure 3.6 (Appendix Figure 3b)** corresponds to the first control sample from skeletal muscle.



Appendix Figure 3 Full lane Western blot images for a) **Figure 3.5** and b) **Figure 3.6**. Each lane corresponds to the first control sample of their respective tissue. Since all Western blots were targeting histone H3 or modified histone H3 residues all lanes are aligned at 17 kDa as indicated by the arrow

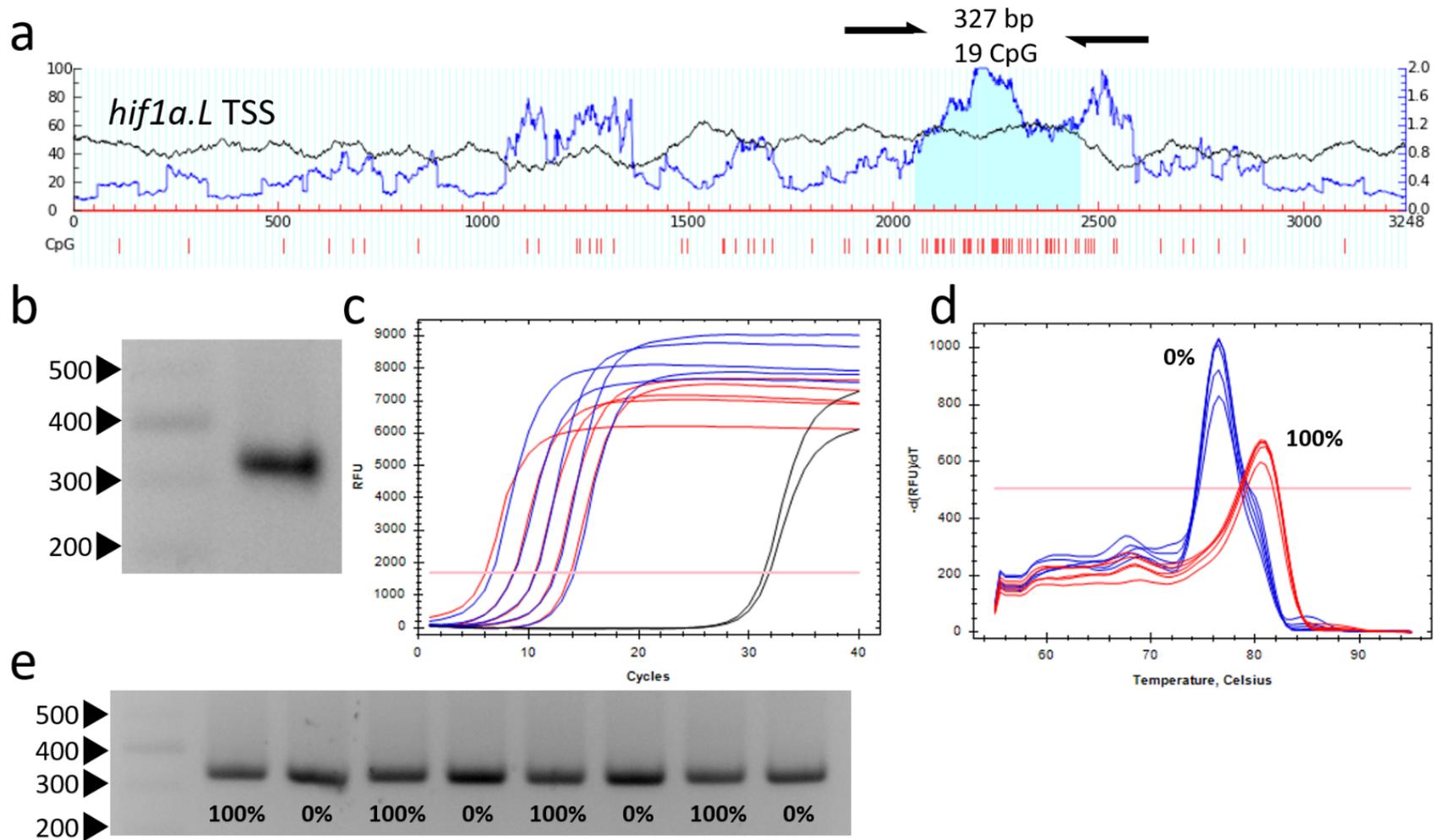
Appendix E: Methylation-sensitive high-resolution melt analysis

MS-HRMA relies on the fact that different sequences of DNA melt at different temperatures. Modern qPCR machines, when utilizing a fully saturable ddDNA dye (EvaGreen), are capable of distinguishing single nucleotide polymorphisms with this method (Liew et al. 2004). DNA that is methylated still has the same sequence of DNA, thus the methylation information must be embedded in the DNA sequence for it to be measured. This is achieved with bisulfite conversion, where unmethylated cytosines are chemically converted to uracil (and subsequently to thymine after PCR), whereas methylated cytosines are protected and remain as cytosines. After bisulfite conversion, differential methylation between two sequences can then be determined since they no longer have the same sequence. To determine how much a sequence is methylated it needs to be compared to a standard curve with varying levels of methylation ranging from fully unmethylated (0%) to fully methylated (100%). To create the standard curve a region of interest (ROI) is selected, for example, upstream of the *hif1 α .L* TSS is a CpG island that can be amplified with methylation naïve primers (**Appendix Figure 4a**). Methylation naïve primers must be able to amplify the same region regardless of methylation status after bisulfite conversion. To start this region is amplified without bisulfite conversion. The product is verified for size and purity on a gel (**Appendix Figure 4b**) and is then purified. After purification the product is split into two samples, one is *in vitro* methylated using SssI CpG methyltransferase (Biotium, Fremont CA) and the other is mock treated. The samples (100% and 0% methylated) are then bisulfite converted and amplified using the methylation naïve primers after serial dilution to ensure no amplification bias is present (**Appendix Figure 4c**). If the *in vitro* methylation was successful there will be a difference in melt temperatures between the 0% and 100%

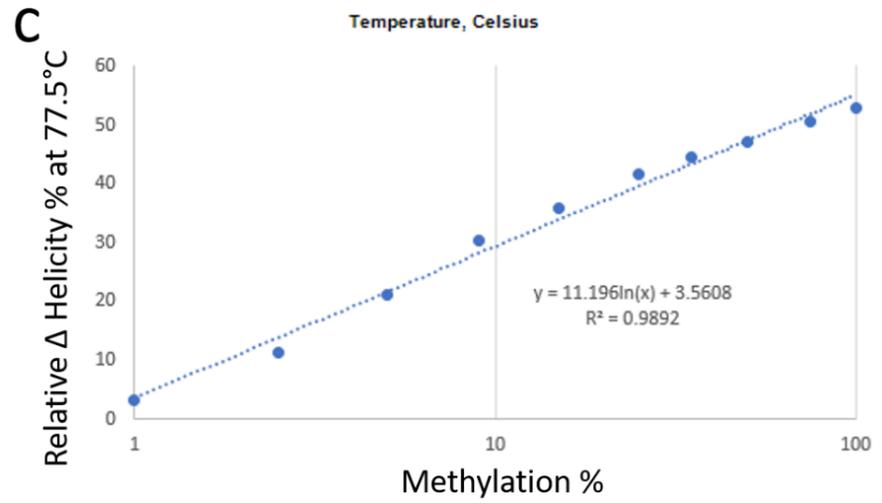
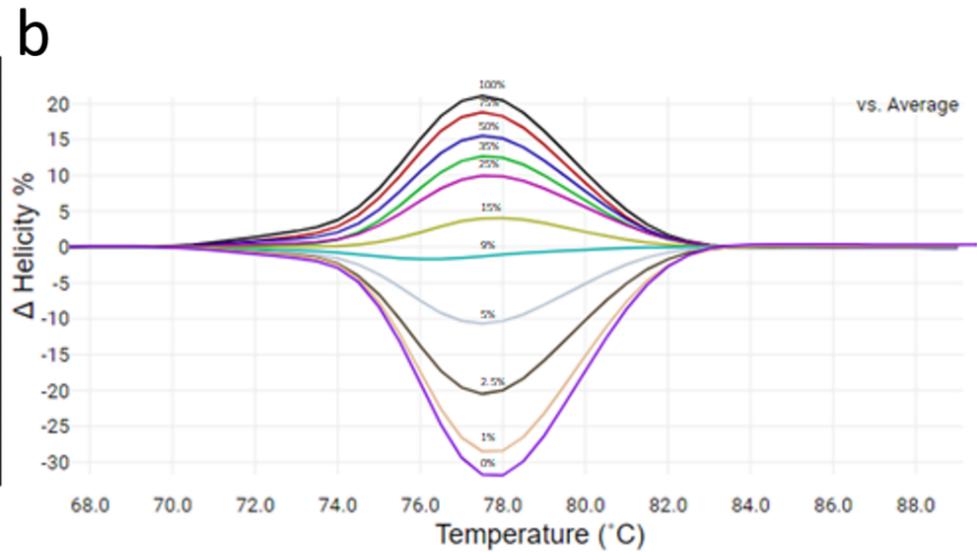
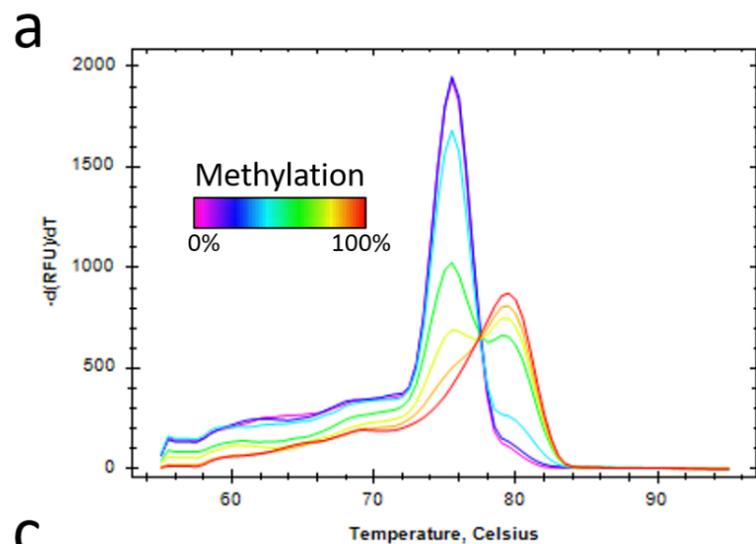
methylated products (**Appendix Figure 4d**), but not a difference in primer size (**Appendix Figure 4e**). This proof of concept can then be scaled to incorporate a range of methylation percentages by mixing the 0% and 100% methylated products together in different proportions (**Appendix Figure 5a**). The melt data can then be normalized using *uAnalyze* (Dwight et al. 2012) (**Appendix Figure 5b**) and a temperature which shows the greatest magnitude changes in helicity is chosen to make the standard curve (**Appendix Figure 5c**). In this case the standard curve covers a broad range of methylation percentages (0, 1, 2.5, 5, 9, 15, 25, 35, 50, 75, 100) and has an R^2 of 0.989. Sample of unknown methylation percentage can then be measured by bisulfite converting the gDNA, amplifying with the methylation naïve primers, normalizing the melt profile with *uAnalyze* and comparing to the standard curve.

References

- Dwight ZL, Palais R, Wittwer CT (2012) *uAnalyze*: web-based high-resolution DNA melting analysis with comparison to thermodynamic predictions. *IEEE/ACM Trans Comput Biol Bioinforma* 9:1805–11 . doi: 10.1109/TCBB.2012.112
- Liew M, Pryor R, Palais R, et al (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem* 50:1156–1164 . doi: 10.1373/clinchem.2004.032136



Appendix Figure 4 Methylation standard curve proof of concept for *hif1a.L* TSS CpG island. a) The region of interest (ROI) was identified near the *hif1a.L* TSS. b) The ROI was amplified and the product was purified. The purified product was then split into two samples, one being *in vitro* methylated by SssI CpG methyltransferase (100% methylation) and the other being mock treated (0%). c) The 0% (blue) and 100% (red) methylated products are then amplified after serial dilution using the same methylation naïve primers to verify no amplification bias, then d) verified to have different melt peaks but e) are still the same size. Results in c), d), and e) indicate that the 0% and 100% methylated products only differ by sequence introduced by the *in vitro* methylation since they are amplified by the same primers and are the same size.



Appendix Figure 5 Full methylation standard curve for *hif1a.L* TSS CpG island. a) A range of methylation percentages is produced by mixing the 0% and 100% methylated product in different proportions. b) The melt profile is normalized using *uAnalyze*. c) The change in helicity at a specific temperature is chosen for the standard curve.